

**New nonrenormalization theorem from UV/IR mixing**Steven Abel<sup>1,\*</sup>, Keith R. Dienes<sup>2,3,†</sup> and Luca A. Nutricati<sup>4,5,‡</sup><sup>1</sup>*Institute for Particle Physics Phenomenology, Durham University, DH1 3LE, Durham, United Kingdom*<sup>2</sup>*Department of Physics, University of Arizona, Tucson, Arizona 85721 USA*<sup>3</sup>*Department of Physics, University of Maryland, College Park, Maryland 20742 USA*<sup>4</sup>*Department of Mathematical Sciences, Durham University, Durham, DH1 3LE United Kingdom*<sup>5</sup>*University College London, Gower Street, London WC1E 6BT United Kingdom*

(Received 26 August 2024; accepted 5 September 2024; published 20 December 2024)

In this paper, we prove a new nonrenormalization theorem which arises from UV/IR mixing. This theorem and its corollaries are relevant for all four-dimensional perturbative tachyon-free closed string theories which can be realized from higher-dimensional theories via geometric compactifications. As such, our theorem therefore holds regardless of the presence or absence of spacetime supersymmetry and regardless of the gauge symmetries or matter content involved. This theorem resolves a hidden clash between modular invariance and the process of decompactification, and enables us to uncover a number of surprising phenomenological properties of these theories. Chief among these is the fact that certain physical quantities within such theories cannot exhibit logarithmic or power-law running and instead enter an effective fixed-point regime above the compactification scale. This cessation of running occurs as the result of the UV/IR mixing inherent in the theory. These effects apply not only for gauge couplings but also for the Higgs mass and other quantities of phenomenological interest, thereby eliminating the logarithmic and/or power-law running that might have otherwise appeared for such quantities. These results illustrate the power of UV/IR mixing to tame divergences—even without supersymmetry—and reinforce the notion that UV/IR mixing may play a vital role in resolving hierarchy problems without supersymmetry.

DOI: [10.1103/PhysRevD.110.126021](https://doi.org/10.1103/PhysRevD.110.126021)**I. INTRODUCTION AND OVERVIEW OF RESULTS**

Nonrenormalization theorems are powerful tools for the study of quantum field theories. Historically, the most famous nonrenormalization theorems are those that arise within the context of theories with unbroken spacetime supersymmetry (SUSY). Such supersymmetric nonrenormalization theorems can be understood as the result of relatively straightforward “level-by-level” pairwise cancellations between states with similar masses, with the renormalization contributions from each particle in the spectrum cancelling against the contributions of corresponding superpartners. However, given that unbroken supersymmetry does not appear to be a feature of the natural world, such SUSY-based nonrenormalization theorems cannot be exact in any phenomenologically viable model. Historically, another (somewhat related) motivation

for SUSY was to address the hierarchy problems of the Standard Model, such as the gauge (Higgs) hierarchy and the cosmological-constant problems. However, given the collider data that has been collected over the past decade, it is also becoming increasingly unlikely that electroweak-scale SUSY plays a significant role in addressing these hierarchy problems.

Recently, increasing attention has focused on the extent to which hierarchy problems can be alternatively addressed through symmetries that involve UV/IR mixing (see, e.g., Refs. [1–17]; for recent reviews see Refs. [18,19]). Within such scenarios, UV physics is directly related to IR physics as well as to physics at intermediate scales, implying that potential solutions to hierarchy problems within such theories might emerge through what might initially appear to be conspiracies across many or all energy scales. However, given the ongoing speculation about the possible existence of such UV/IR-mixed approaches to hierarchy problems, we are then led to ask the further question as to whether such UV/IR-mixed symmetries might also give rise to nonrenormalization theorems. Such nonrenormalization theorems would then emerge not as the consequence of pairwise level-by-level cancellations (such as those that arise in theories with unbroken supersymmetry), but instead as the consequence of symmetries which operate across all scales simultaneously.

\*Contact author: [s.a.abel@durham.ac.uk](mailto:s.a.abel@durham.ac.uk)†Contact author: [dienes@arizona.edu](mailto:dienes@arizona.edu)‡Contact author: [l.nutricati@ucl.ac.uk](mailto:l.nutricati@ucl.ac.uk)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

In this paper, we investigate this issue by focusing on one of the most important UV/IR-mixed symmetries in string theory, specifically *worldsheet modular invariance*. Worldsheet modular invariance is an exact fundamental symmetry within perturbative closed string theories, and remnants of modular invariance even exist for open strings as well. Within a given string model, modular invariance governs the string spectrum and its interactions regardless of the presence or absence of spacetime supersymmetry, regardless of the gauge group and particle content of the model, and even regardless of its assumed spacetime dimensionality. While modular invariance has numerous implications for the low-energy phenomenologies of such strings, this paper is devoted to demonstrating that worldsheet modular invariance has an additional effect which has not previously been noticed, namely that it also gives rise to a powerful nonrenormalization theorem.

As we shall find, this theorem emerges within the context of perturbative four-dimensional tachyon-free closed string theories which can be realized as geometric compactifications of higher-dimensional string theories. In other words, our theorem applies to all perturbative four-dimensional closed string theories which have self-consistent decompactification limits. This gives our theorem a broad and nearly universal applicability.

Rigorously stating and proving this theorem is one of the primary goals of this paper. However, in order to gain a very rough understanding of the content of our theorem, we shall here provide an overview of our main results. To do this, let us begin by recalling that as the spacetime dimensionality of an ordinary quantum field theory increases, it tends to become more finite in the IR but more divergent in the UV. This is the direct result of the different numbers of momentum components associated with the states propagating in loops. However, in theories with UV/IR mixing, it turns out that this behavior is generally different. For example, within closed string theories there is only one potential divergence for a given one-loop amplitude. Indeed, by making use of the UV/IR-mixed symmetries of the theory, one can recast this divergence as either a UV divergence or an IR divergence [13]. Moreover, as the dimensionality of such a string theory increases, it turns out that the theory tends to become *more* finite (or equivalently, its divergences tend to become less severe). This feature arises because additional internal cancellations or constraints come into play across the string spectrum as the spacetime dimensionality of our string theory increases—constraints which soften or eliminate these divergences and which are thereby ultimately responsible for these additional finiteness properties [1,2].

This situation becomes particularly interesting for string theories which have decompactification limits. If a given string theory has a *bona fide* decompactification limit, then its spectrum must satisfy the extra constraints discussed above in the full decompactification limit at which our

theory becomes higher-dimensional. By contrast, there is no need for such extra constraints to apply to the compactified theory, since the compactified theory by definition is lower-dimensional. What we find, however, is that these extra constraints apply not only to the higher-dimensional theory that emerges in the full decompactification limit, *but also* to the original compactified theory, regardless of the compactification volume. In particular, stated succinctly, in this paper we shall prove the following:

*Theorem.* Any four-dimensional closed string theory which can be realized as a geometric compactification from a higher-dimensional string theory will inherit the precise stricter internal cancellations of the higher-dimensional theory from which it is obtained despite the compactification.

Thus, as long as our four-dimensional theory has a decompactification limit, its spectrum must satisfy not only the constraints that would normally be associated with its existence in four dimensions, but also the additional constraints that would be required in higher dimensions. Indeed, this remains true even if our four-dimensional theory is nowhere near the decompactification limit and is thus fully four-dimensional. As we shall demonstrate, this theorem and the cancellations it requires are realized across all energy scales at once, and the mechanism by which it operates has no field-theoretic analog or approximation.

This theorem leads to many surprising phenomenological consequences for our original four-dimensional theory. One of these consequences is that there are new, unexpected UV/IR-mixed supertrace constraints which operate across the entire four-dimensional string spectrum at all energy scales, similar to those which were originally obtained in Ref. [2] and more recently generalized in Refs. [13,17]. Like these previous supertrace constraints, our new supertrace constraints are the results of an underlying so-called “misaligned supersymmetry” [1,2] that governs the spectra of all tachyon-free modular-invariant theories—even those that lack spacetime supersymmetry. We emphasize that these new constraints are completely unexpected from the perspective of our original four-dimensional theory. They nevertheless secretly govern the spectra of such theories at all energy scales.

Another surprising conclusion of our theorem concerns the effective field theories (EFTs) that are derived from such string theories. In particular, within such theories we find that the couplings can at most run *only until the compactification scale is reached*. Beyond this point, our theorem asserts that *all running ceases*—even if this compactification scale is much lower than the string scale. Indeed, the theory necessarily enters a “fixed-point” regime in which the beta-functions of the gauge couplings vanish. This too is a result of the extra “hidden” constraints discussed above. This gives us an important corollary to our main theorem:

*Nonrenormalization corollary.* Within any modular-invariant theory which has  $\delta \equiv D - 4$  large extra dimensions

opening up at a scale  $1/R$ , misaligned supersymmetry and UV/IR mixing eliminate all running for  $\mu \gtrsim R^{-1}$  regardless of the value of  $\delta$ . For  $\mu < R^{-1}$ , these same phenomena eliminate all running for  $\delta > 2$ , and leave at most logarithmic running for  $\delta = 2$ .

The above results arise most naturally within the context of string theories exhibiting a single decompactification limit (along with a corresponding  $T$ -dual limit). However, most string theories have *multiple* distinct decompactification limits, and the different higher-dimensional theories which emerge in these limits may even have different spacetime dimensionalities. Such cases nevertheless continue to satisfy our theorem. In particular, we shall find that four-dimensional theories with multiple decompactification limits will simultaneously satisfy *all* of the different constraints that emerge from each individual decompactification. Moreover, while our discussion here will focus on the case of four-dimensional theories, there is nothing intrinsically special to four dimensions, and similar results apply for theories in other spacetime dimensions as well, as long as such theories continue to have their own decompactification limits.

Along the way, we also prove another potentially important result. Specifically, we prove that the one-loop contributions to certain string amplitudes have a universal behavior in the limit of large compactification volume. In particular, we define a modular-invariant compactification volume  $\tilde{V}_T$ , and then demonstrate that all such amplitudes necessarily scale as  $(\pi/3)\tilde{V}_T$  as  $\tilde{V}_T \rightarrow \infty$ .

In this Introduction we have merely sought to describe our theorem and how it operates in a rough intuitive sense. Needless to say, there are many subtle details which we are omitting here. These details will be discussed in subsequent sections. Moreover, as might be imagined, this paper is somewhat technical and relies on a number of results which were established in previous papers by the present authors and others. We have therefore attempted to keep this paper entirely self-contained by including an initial section in which we summarize those aspects of previous work which will ultimately be relevant for proving our theorem.

Accordingly, this paper is organized as follows. First, in Sec. II, we assemble all of the conceptual and calculational ingredients that will ultimately be needed for proving our theorem. Then, in Sec. III, we discuss a fundamental clash that emerges between modular invariance and the process of decompactification, and explain how our theorem automatically resolves this clash. Thus Sec. IV may be regarded as the central crux of this paper in which we present our theorem and discuss how it fits into (and emerges from) the larger theoretical framework. In Sec. IV, we then proceed to discuss two of the most important implications of our theorem. These include not only new supertrace constraints to which our theorem gives rise, but also tight restrictions on the running of couplings in these UV/IR-mixed theories. Finally, for the experts, in Sec. V

we provide an explicit example in which all of these results are illustrated through direct calculation. We then conclude in Sec. VI with a discussion of some of the additional consequences of our theorem for low-energy string phenomenology.

## II. ASSEMBLING THE INGREDIENTS

In this section we collect the central ingredients that will be required in order to formulate, prove, and interpret our nonrenormalization theorem.

### A. Operator insertions

In general, a given perturbative closed string theory formulated in  $D$  uncompactified spacetime dimensions will have a partition function  $Z^{(D)}$  which is a function of a modular parameter  $\tau \equiv \tau_1 + i\tau_2$  with  $\tau_i \in \mathbb{R}$ , and which can be written as a double power-series expansion in  $q \equiv e^{2\pi i\tau}$  and  $\bar{q} \equiv e^{-2\pi i\bar{\tau}}$  of the form

$$Z^{(D)}(\tau) = \tau_2^k \sum_{m,n} a_{mn} \bar{q}^m q^n. \quad (2.1)$$

Here the summation is over all right-moving and left-moving worldsheet energy levels of the string, respectively denoted  $(m, n)$ , and  $a_{mn}$  is the net (bosonic minus fermionic) number of degrees of freedom in the string spectrum with worldsheet energies  $(m, n)$ . Physical consistency of the partition function requires that it be modular invariant, i.e., that  $Z(\tau + 1) = Z(-1/\tau) = Z(\tau)$ . It is the latter invariance under  $\tau \rightarrow -1/\tau$  which ties together the degeneracies of states  $a_{mn}$  at all energies  $(m, n)$  across the string spectrum, thereby connecting UV and IR physics in a highly nontrivial way. The quantity  $k$  is the *modular weight* of the partition function, and for a string theory formulated in  $D$  uncompactified spacetime dimensions generically has the value

$$k = 1 - D/2. \quad (2.2)$$

We thus have  $k = -1$  for  $D = 4$ . In general, the space-time mass  $M$  of any state with worldsheet energies  $(m, n)$  is given by

$$M^2 = \frac{1}{2}(M_R^2 + M_L^2) = \frac{2}{\alpha'}(m + n), \quad (2.3)$$

where  $m = \alpha' M_R^2/4$ ,  $n = \alpha' M_L^2/4$ , and  $\alpha' \equiv 1/M_s^2$ , where  $M_s$  is the string scale. The states with  $m = n$  are considered “on-shell” or “physical” and can serve as in- and out-states, while the “off-shell” states with  $m \neq n$  are intrinsically stringy or “unphysical” and contribute only in loop amplitudes. We shall assume throughout this paper that we are dealing with *tachyon-free* string theories (i.e., theories for which  $a_{nn} = 0$  for all  $n < 0$ ). However, we shall not make any assumption that our theory exhibits

spacetime supersymmetry. Thus we will not assume that  $a_{nn} = 0$  for all non-negative values of  $n$ , or make any other similar assumptions regarding the vanishing of the  $a_{mn}$  coefficients beyond our tachyon-free requirement that  $a_{nn} = 0$  for all  $n < 0$ . In this connection, we note that no phenomenologically viable model can remain exactly supersymmetric. By contrast, all string models must maintain an exact modular invariance as part of their internal self-consistency constraints.

In this paper, we consider the corresponding one-loop amplitudes associated with various physical quantities in four dimensions. In particular, we focus on amplitudes in which there are no (or small) external momenta, or alternatively amplitudes in which such external momenta can be factored out from the one-loop integration. This is a large and crucial class of amplitudes, and we shall see explicit examples below. We shall let  $\zeta$  generally denote the phenomenological quantities for which such amplitudes provide the one-loop contributions.

In general we begin the calculation of such one-loop amplitudes by calculating the modular-invariant tally of the contributions to  $\zeta$  coming from each string state. With the assumptions described above, this tally will take the form

$$Z_{\mathcal{X}}^{(4)} = \tau_2^{-1} \sum_{m,n} a_{mn} \mathcal{X}_{mn} \bar{q}^m q^n. \quad (2.4)$$

This clearly resembles the partition function  $Z^{(4)}$  but also includes a factor  $\mathcal{X}_{mn}$  which denotes the contribution to  $\zeta$  from each  $(m, n)$  state. The resulting one-loop contribution  $\zeta$  is then given by

$$\zeta \equiv \langle \mathcal{X} \rangle \quad (2.5)$$

where

$$\langle \mathcal{X} \rangle \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z_{\mathcal{X}}^{(4)}. \quad (2.6)$$

Here  $d^2\tau/\tau_2^2$  is the standard modular-invariant measure and  $\mathcal{F}$  is the fundamental domain of the modular group

$$\mathcal{F} \equiv \{\tau: |\tau_1| \leq 1/2, |\tau| \geq 1, \tau_2 > 0\} \quad (2.7)$$

with  $\tau_1 \equiv \text{Re } \tau$  and  $\tau_2 \equiv \text{Im } \tau$  respectively.

In general, these factors  $\mathcal{X}_{mn}$  within Eq. (2.4) are the eigenvalues of an *operator insertion*  $\mathcal{X}$  into the partition function. In general, there is a different operator insertion  $\mathcal{X}$  for each physical quantity  $\zeta$ . In this paper we concentrate on operator insertions  $\mathcal{X}$  which take the form

$$\mathcal{X} = \mathbb{X}_0 + \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \quad (2.8)$$

where each  $\mathbb{X}_\ell$  is  $\tau$ -independent. However it turns out that the operator insertions for any physical quantity in four

dimensions either take the form in Eq. (2.8) directly or can be reduced to it. Thus, we may consider the operator-insertion form in Eq. (2.8) to be completely general.

In general, just like the partition function  $Z^{(4)}$  itself, the resulting  $\mathcal{X}$ -weighted spectral tally  $Z_{\mathcal{X}}^{(4)}$  must also be modular invariant. This in turn implies that  $\mathcal{X}$  must be a modular-invariant operator insertion, which further implies that the  $\mathbb{X}_\ell$ -coefficients in Eq. (2.8) are tightly linked together by modular invariance. Thus, knowledge of any one of these  $\mathcal{X}_\ell$ -insertions permits the determination of the others through a process of modular completion [13]. In all cases, however, the requirements of modular invariance ensure that  $\mathbb{X}_0$  can be at most proportional to the identity operator 1. Thus the one-loop contribution to  $\zeta$  from  $\mathbb{X}_0$  will be proportional to the four-dimensional one-loop cosmological constant

$$\Lambda \equiv -\frac{\mathcal{M}^4}{2} \langle \mathbf{1} \rangle = -\frac{\mathcal{M}^4}{2} \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z^{(4)} \quad (2.9)$$

where  $\mathcal{M}$  is the reduced string scale  $M_s/(2\pi)$ .

Later in Sec. IV and the Appendix we will extend our analysis to certain cases in which the  $\mathbb{X}_\ell$  can carry a holomorphic dependence on  $\tau$ . We shall find, however, that such cases do not disturb our main results.

As we shall see, our theorem and its proof will not require any further details regarding these operator insertions  $\mathbb{X}_\ell$ . However, it may be useful to recall what these insertions look like in various phenomenologically relevant cases. As a first example, we may consider  $\zeta$  to be the one-loop contribution to the mass of a scalar Higgs field in an arbitrary heterotic string model. This calculation is discussed in detail in Ref. [13], where it shown that the corresponding operator-insertion coefficients  $\mathbb{X}_\ell$  turn out to be given by

$$\begin{aligned} \mathbb{X}_0 &\equiv -\frac{\xi \mathcal{M}^2}{8\pi^2} \\ \mathbb{X}_1 &\equiv \frac{\mathcal{M}^2}{8\pi} \partial_\phi^2 M^2 \Big|_{\phi=0} \\ \mathbb{X}_2 &\equiv -\frac{1}{32\pi^2} (\partial_\phi M^2)^2 \Big|_{\phi=0} \end{aligned} \quad (2.10)$$

where  $M(\phi)$  describes the mass of a given string state as a function of the fluctuation  $\phi$  in the particular Higgs field in question. Thus, the functions  $M(\phi)$ —and whether they are zero or not—essentially encode which Higgs field  $\phi$  is under discussion. The parameter  $\xi$  is a function of the shifts induced on the mass spectrum by the Higgs field.

To make these expressions for  $\mathbb{X}_\ell$  more explicit, we may reexpress them in terms of charge operators  $\mathbf{Q}$  which fill out the so-called ‘‘charge lattice’’ associated with the string spectrum [13]. In general, these charge operators take the form  $\mathbf{Q} = (\mathbf{Q}_L, \mathbf{Q}_R)$  where the ‘L’ and ‘R’ components

correspond to the charges associated with our left-moving and right-moving worldsheet degrees of freedom. For perturbative heterotic strings in four spacetime dimensions, these lattices have maximal dimensionalities (22,10); the left-moving charge components are generally associated with the corresponding gauge group while the right-moving components generally correspond to additional factors such as spacetime helicities. In general, these charge lattices are also *Lorentzian*, meaning that the scalar dot-product between two charge operators is defined as  $\mathbf{Q} \cdot \mathbf{Q}' \equiv \mathbf{Q}_L \mathbf{Q}'_L - \mathbf{Q}_R \mathbf{Q}'_R$ . In terms of these lattice operators, the action of a nonzero Higgs VEV  $\phi$  is to induce a shift in the lattice of  $\mathbf{Q}$  charges [13]:

$$\mathbf{Q} \rightarrow \mathbf{Q} + \frac{\phi}{M_s} \mathcal{T} \cdot \mathbf{Q} + \mathcal{O}(\phi^2) \quad (2.11)$$

where the “response” matrix  $\mathcal{T}$  can be decomposed in a left-right block-diagonal fashion as

$$\mathcal{T} = \begin{pmatrix} \mathcal{T}_{LL} & \mathcal{T}_{LR} \\ -\mathcal{T}_{LR}^t & \mathcal{T}_{RR} \end{pmatrix}. \quad (2.12)$$

This then yields alternative expressions for the  $\mathbb{X}_1$  and  $\mathbb{X}_2$  operators in Eq. (2.10) as sums over charges:

$$\begin{aligned} \mathbb{X}_1 &= \frac{\mathcal{M}^2}{2\pi} (\mathbf{Q} \cdot \mathcal{S} \cdot \mathbf{Q}) \\ \mathbb{X}_2 &= -\mathcal{M}^2 (\mathbf{Q} \cdot \mathcal{T} \cdot \mathbf{Q})^2 \end{aligned} \quad (2.13)$$

where we have defined

$$\mathcal{S} = \begin{pmatrix} \mathcal{T}_{LR} \mathcal{T}_{LR}^t & -\mathcal{T}_{LL} \mathcal{T}_{LR} \\ \mathcal{T}_{RR} \mathcal{T}_{LR}^t & \mathcal{T}_{LR}^t \mathcal{T}_{LR} \end{pmatrix}. \quad (2.14)$$

Meanwhile  $\mathbb{X}_0$  is given by [13]

$$\xi \equiv \frac{1}{2} \text{Tr} \mathcal{S}, \quad (2.15)$$

whereupon

$$\mathbb{X}_0 \equiv -\frac{\mathcal{M}^2}{16\pi^2} \text{Tr} \mathcal{S}. \quad (2.16)$$

In a similar fashion, we can also consider the case in which  $\zeta$  is related to the one-loop contribution to the gauge coupling  $g_G$  for any group factor  $G$  in a given string model. This case is discussed in detail in Ref. [17]. To perform this calculation, we evaluate these couplings  $g_G$  to one-loop order and then separate out the tree-level contributions. In general, these quantities are related through

$$\left. \frac{16\pi^2}{g_G^2} \right|_{\text{total}} = \left. \frac{16\pi^2}{g_G^2} \right|_{\text{tree}} + \Delta_G \quad (2.17)$$

where in string theory we have  $g_G|_{\text{tree}} \sim e^{-\langle \phi \rangle}$  with  $\langle \phi \rangle$  denoting the VEV of the dilaton  $\phi$  and where  $\Delta_G$  denotes the one-loop contribution to  $16\pi^2/g_G^2$ . We may thus now take  $\zeta$  to be the one-loop contribution  $\Delta_G$ , whereupon the corresponding operator insertions are given by [17]

$$\begin{aligned} \mathbb{X}_0 &= 0 \\ \mathbb{X}_1 &\equiv \frac{\xi}{2\pi} \left( \bar{Q}_H^2 - \frac{\bar{E}_2}{12} \right) \\ \mathbb{X}_2 &\equiv -2 \left( \bar{Q}_H^2 - \frac{\bar{E}_2}{12} \right) Q_G^2 \end{aligned} \quad (2.18)$$

where  $\bar{Q}_H$  is the (right-moving) spacetime helicity operator (a specific component of  $\mathbf{Q}_R$ ) and where  $Q_G^2$  is the quadratic Casimir of  $G$  (comprised out of components of  $\mathbf{Q}_L$ ). In Eq. (2.18), the quantity  $\bar{E}_2$  is the antiholomorphic Eisenstein function which is defined in Eq. (A3). As discussed above, this case therefore provides an example in which the  $\mathbb{X}_\ell$  are not  $\tau$ -independent but rather carry a holomorphic  $\tau$ -dependence. Such cases will be considered in Sec. IV, but we shall find that they will not induce significant departures from the main results we shall be presenting.

## B. Divergences and regulator function

In general, with operator insertions  $\mathcal{X}$  of the form in Eq. (2.8), it is possible that the four-dimensional amplitude in Eq. (2.6) experiences a logarithmic divergence. Indeed, such a divergence will arise in four-dimensional theories if

$$\text{Str}_{M=0} \mathbb{X}_2 \neq 0 \quad (2.19)$$

where  $\text{Str}_{M=0}$  denotes a supertrace over only the massless states. Given that our operator insertions generally take the form in Eq. (2.8), this is the most severe divergence that can arise.

In such cases, we must regulate the amplitude without disturbing its modular invariance. Following Refs. [13,17] we will carry out this procedure by multiplying the integrand of Eq. (2.6) by a suitable modular-invariant regulator function  $\mathcal{G}(a, \tau)$  where  $\tau$  is the one-loop modular parameter and  $a$  schematically represents other possible parameters within this function. In order to serve its purpose as a regulator, such a function must vanish more rapidly than logarithmically as  $\tau \rightarrow i\infty$ . We also demand that  $\mathcal{G} \approx 1$  elsewhere within the fundamental domain, so that this regulator does not significantly disturb our theory within regions of integration that do not lead to divergences.

An explicit regulator function satisfying these criteria was given in Ref. [13]. However, given that the specific form of this function will not be needed for any of our main arguments, we shall defer discussion of this function until

Sec. IV, when we work out a specific example of our results.

Thus, our procedure for regulating a divergent one-loop string amplitude amounts to deforming the amplitude according to our regulator function  $\mathcal{G}$ :

$$\langle \mathcal{X} \rangle \rightarrow \langle \mathcal{X} \rangle_{\mathcal{G}} \equiv \langle \mathcal{X}\mathcal{G} \rangle. \quad (2.20)$$

As we shall see, there can also be other reasons for introducing such a regulator function, even for amplitudes that are in principle finite.

### C. Rankin-Selberg transform and supertraces over physical string states

In general, for arbitrary insertion  $\mathcal{X}$  and arbitrary dimension  $D$ , the one-loop amplitude  $\langle \mathcal{X} \rangle$  is given by the  $D$ -dimensional version of Eq. (2.6), namely

$$\langle \mathcal{X} \rangle \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z_{\mathcal{X}}^{(D)}. \quad (2.21)$$

We thus see that string states with all allowed combinations of worldsheet energies  $(m, n)$  contribute. This is true not only for the on-shell states with  $m = n$  but also the off-shell states with  $m \neq n$ ; indeed, while the former contribute at all values of  $\tau_2$  within the fundamental domain in Eq. (2.7), the latter also contribute, albeit within only the  $\tau_2 < 1$  region. These contributions can nevertheless be sizable.

It turns out that this amplitude may be expressed in another form which depends only on the *on-shell* states with  $m = n$ . This alternate form will be critical for our eventual theorem, and exists for all dimensions  $D$  and for all situations in which the amplitude  $\langle \mathcal{X} \rangle$  is finite. Specifically, as long as the amplitude  $\langle \mathcal{X} \rangle$  in Eq. (2.21) is finite and modular invariant, a powerful result in modular-function theory due to Rankin [20,21] and Selberg [22] allows us to reexpress this amplitude as

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \operatorname{Res}_{s=1} \int_0^{\infty} d\tau_2 \tau_2^{s-2} g(\tau_2) \quad (2.22)$$

where

$$\begin{aligned} g(\tau_2) &\equiv \int_{-1/2}^{1/2} d\tau_1 Z_{\mathcal{X}}^{(D)}(\tau) \\ &\equiv \tau_2^k \int_{-1/2}^{1/2} d\tau_1 \sum_{m,n} a_{mn} \mathcal{X}_{mn} \bar{q}^m q^n \\ &= \tau_2^k \sum_n a_{nn} \mathcal{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \end{aligned} \quad (2.23)$$

with  $\alpha' M_n^2 = 4n$ , as in Eq. (2.3). This tells us that the original string amplitude  $\langle \mathcal{X} \rangle$  is nothing but the Mellin transform of  $g(\tau_2)/\tau_2$ . This in turn allows us to write  $g(\tau_2)$

directly as the inverse Mellin transform of the amplitude, which yields the alternative result [23,24]

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g(\tau_2), \quad (2.24)$$

where  $g(\tau_2)$  continues to be given by Eq. (2.23). This result is equivalent to that in Eq. (2.22), but has the primary advantage that we can now evaluate  $\langle \mathcal{X} \rangle$  simply by taking the  $\tau_2 \rightarrow 0$  limit of  $g(\tau_2)$  rather than by evaluating the residue of the  $\tau_2$ -integral of  $g(\tau_2)$ . Indeed, inserting Eq. (2.23) into Eq. (2.24) yields

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^k \sum_n a_{nn} \mathcal{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2}. \quad (2.25)$$

This, then, expresses the amplitude  $\langle \mathcal{X} \rangle$  in terms of the degeneracies  $a_{nn}$  of only the *physical* string states.

The issue then boils down to the evaluation of the right side of Eq. (2.25). Following Ref. [17], we shall do this by first defining the sum

$$S^{(D)}(\tau_2) \equiv \tau_2^{-k} g(\tau_2) = \sum_n a_{nn} \mathcal{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2}. \quad (2.26)$$

Thus  $S^{(D)}$  encapsulates only that part of the  $\tau_2$ -dependence of  $g(\tau_2)$  that comes from the  $\mathcal{X}$ -weighted sums over the string states. We can then expand  $S^{(D)}(\tau_2)$  as a power series in  $\tau_2$ , i.e.,

$$S^{(D)}(\tau_2) \sim \sum_j C_j \tau_2^j \quad \text{as } \tau_2 \rightarrow 0. \quad (2.27)$$

Note that for complete generality we will not assume that only integer values of  $j$  contribute in Eq. (2.27); indeed subleading contributions can generically also have fractional  $j$ . Inserting Eq. (2.27) into Eq. (2.24), we thus have

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \sum_j C_j \tau_2^{j+k}. \quad (2.28)$$

It is not difficult to determine the values of the  $C$ -coefficients for integer  $j$ . Indeed, following Ref. [2], we may “invert” Eq. (2.27) by taking  $\tau_2$ -derivatives of both sides. In this way we find

$$C_0 = \lim_{\tau_2 \rightarrow 0} \left[ \sum_n a_{nn} \mathcal{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \right] \quad (2.29)$$

and

$$\begin{aligned}
 C_1 &= \lim_{\tau_2 \rightarrow 0} \left\{ \sum_n a_{nn} \left[ \frac{d}{d\tau_2} \mathcal{X}_{nn} \right] e^{-\pi\alpha' M_n^2 \tau_2} \right\} \\
 &\quad - \lim_{\tau_2 \rightarrow 0} \left[ \sum_n a_{nn} \mathcal{X}_{nn} (\pi\alpha' M_n^2) e^{-\pi\alpha' M_n^2 \tau_2} \right] \\
 &= \lim_{\tau_2 \rightarrow 0} \left[ \sum_n a_{nn} \left( \frac{d}{d\tau_2} - \pi\alpha' M_n^2 \right) \mathcal{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \right] \\
 &= \lim_{\tau_2 \rightarrow 0} \left[ \sum_n a_{nn} (D_{\tau_2} \mathcal{X}_{nn}) e^{-\pi\alpha' M_n^2 \tau_2} \right] \quad (2.30)
 \end{aligned}$$

where  $D_{\tau_2}$  is the modular-covariant derivative

$$\begin{aligned}
 D_{\tau_2} &\equiv \frac{d}{d\tau_2} - \pi\alpha' M^2 \\
 &= \frac{d}{d\tau_2} - \frac{M^2}{4\pi\mathcal{M}^2}. \quad (2.31)
 \end{aligned}$$

Coefficients  $C_j$  with integer  $j \geq 2$  can be calculated in a similar fashion by taking additional  $\tau_2$ -derivatives, yielding the general result

$$C_j = \frac{1}{j!} \lim_{\tau_2 \rightarrow 0} \left[ \sum_n a_{nn} (D_{\tau_2}^j \mathcal{X}_{nn}) e^{-\pi\alpha' M_n^2 \tau_2} \right]. \quad (2.32)$$

These results may be further simplified by defining the regulated supertrace [2]

$$\text{Str} \mathcal{A} \equiv \lim_{\tau_2 \rightarrow 0} \sum_n a_{nn} \mathcal{A}_{nn} e^{-\pi\alpha' M_n^2 \tau_2}. \quad (2.33)$$

Given that the  $a_{nn}$  coefficients tally the net number of bosonic minus fermionic degrees of freedom with left- and right-moving worldsheet energies  $n$ , such supertraces are completely analogous to the standard spectral supertraces  $\text{Str} \mathcal{A} \equiv \sum_n (-1)^F A_n$  that we would have in ordinary quantum field theory except that they yield finite results even for UV/IR-mixed theories which contain infinite towers of states [2]. Indeed, even in such cases the summation in Eq. (2.33) is finite thanks to the exponential damping factor  $e^{-\pi\alpha' M_n^2 \tau_2}$ , and remains finite even as  $\tau_2 \rightarrow 0$  and this damping factor is removed. We shall therefore adopt this supertrace definition in what follows. Expressed in terms of these supertraces, we then find that our  $C_j$ -coefficients with integer  $j$  all take the relatively simple form

$$C_j = \frac{1}{j!} \text{Str} D_{\tau_2}^j \mathcal{X} \quad \text{for all } j \geq 0. \quad (2.34)$$

Before proceeding further, we emphasize that the above derivation leading to the supertrace expression for the  $C$ -coefficients in Eq. (2.34) implicitly rested on an understanding that  $\text{Str} \mathcal{A} = 0$  if  $\mathcal{A}$  itself contains an uncancelled

positive power of  $\tau_2$ . This follows formally from the fact that the definition of the supertrace in Eq. (2.33) includes a limit taking  $\tau_2 \rightarrow 0$ . It may indeed seem somewhat unorthodox to have  $\tau_2$  appear not only within the argument of the supertrace but also as the supertrace regulator, but expressions such as that in Eq. (2.34) have a clear operational definition and will cause no difficulty. Thus, for example, if  $\mathcal{X}$  takes the form in Eq. (2.8) with  $\tau_2$ -independent coefficients  $\mathbb{X}_\ell$ , then

$$\begin{aligned}
 \text{Str} \mathcal{X} &= \text{Str} \mathbb{X}_0 \\
 \text{Str} \frac{d}{d\tau_2} \mathcal{X} &= \text{Str} \mathbb{X}_1 \\
 \text{Str} \frac{d^2}{d\tau_2^2} \mathcal{X} &= 2 \text{Str} \mathbb{X}_2. \quad (2.35)
 \end{aligned}$$

The result in Eq. (2.28) enables us to express our string amplitude  $\langle \mathcal{X} \rangle$  in terms of the  $C$ -coefficients. For example, taking  $k = -1$  (as appropriate for string theories in four dimensions) and recalling that  $\langle \mathcal{X} \rangle$  is finite, we find

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} C_1. \quad (2.36)$$

However, this result assumes that we have also imposed the auxiliary conditions

$$C_j = 0 \quad \text{for all } j < 1. \quad (2.37)$$

In particular, from this we learn that

$$C_0 = 0. \quad (2.38)$$

The result in Eq. (2.36) allows us to express our string amplitude  $\langle \mathcal{X} \rangle$  in terms of supertraces over only the physical string states. Indeed, combining Eqs. (2.30) and (2.36) we have

$$\begin{aligned}
 \langle \mathcal{X} \rangle &= \frac{\pi}{3} \text{Str} \left( \frac{d}{d\tau_2} \mathcal{X} \right) - \frac{\pi}{3} \text{Str} [\mathcal{X} (\pi\alpha' M^2)] \\
 &= \frac{\pi}{3} \text{Str} (D_{\tau_2} \mathcal{X}). \quad (2.39)
 \end{aligned}$$

Likewise, our auxiliary condition in Eq. (2.38) now takes the form

$$\text{Str} \mathcal{X} = 0. \quad (2.40)$$

Note that these results apply for any modular-invariant operator insertion  $\mathcal{X}$  in four dimensions so long as this insertion results in a finite amplitude  $\langle \mathcal{X} \rangle$ .

Finally, we note that we may occasionally be required to evaluate supertraces of quantities—such as the  $\mathbb{X}_\ell$  in Eq. (2.18)—which involve the Eisenstein function  $E_2(\tau)$  defined in Eq. (A3). The presence of such a function

introduces a number of subtleties into the procedure for evaluating supertraces. These subtleties are fully described in Ref. [17] and summarized in the Appendix, with the end result that the usual notion of supertrace is generalized in a straightforward manner.

The results in Eqs. (2.36) and (2.38) were derived for  $k = -1$ , as appropriate for four-dimensional theories. However, these results can be easily generalized to higher dimensions  $D$ . Indeed, given the relation in Eq. (2.2), we obtain

$$\begin{aligned}
 D = 6: & \begin{cases} C_0 = C_1 = 0 \\ \langle \mathcal{X} \rangle = \frac{\pi}{3} C_2 \end{cases} \\
 D = 8: & \begin{cases} C_0 = C_1 = C_2 = 0 \\ \langle \mathcal{X} \rangle = \frac{\pi}{3} C_3 \end{cases} \\
 D = 10: & \begin{cases} C_0 = C_1 = C_2 = C_3 = 0 \\ \langle \mathcal{X} \rangle = \frac{\pi}{3} C_4. \end{cases} \quad (2.41)
 \end{aligned}$$

We thus observe, as discussed in the Introduction, that theories in higher dimensions exhibit more internal cancellation constraints than do theories in lower dimensions.

Note that for convenience we shall restrict our attention in this paper to spacetime dimensionalities which are even, with  $D \in 2\mathbb{Z}$ . These are the dimensionalities for which the modular weight  $k$  is an integer, and in which chiral theories can exist. Similar results also arise for odd  $D$ , but with additional complications that obscure the underlying physics. We shall therefore focus on theories with even  $D$  in what follows.

### III. THE THEOREM

Having assembled the ingredients that will be needed for our theorem, we now turn our attention to the theorem itself. Our theorem ultimately rests on modular invariance and misaligned supersymmetry, as do most of the results quoted in Sec. II. As we shall see, there is a deep clash between modular invariance and the process of decompactification. This clash is intrinsic to the UV/IR mixing inherent in string theory, and does not exist in ordinary quantum field theory. Our theorem emerges as the result of this clash, and ultimately provides the means by which these two features can be reconciled.

#### A. Fundamental clash between decompactification and modular invariance

Let us begin by examining the properties of modular-invariant four-dimensional theories in the presence of a large-volume  $\delta$ -dimensional compactification. In other words, we shall consider a  $(4 + \delta)$ -dimensional modular-invariant theory compactified on a manifold of the form  $\mathcal{M}_4 \times \mathcal{K}_\delta$  where  $\mathcal{M}_4$  is ordinary uncompactified Minkowski space and where  $\mathcal{K}_\delta$  is our compactified  $\delta$ -dimensional space whose characteristic dimensions we shall consider to be large, with a

corresponding  $\delta$ -dimensional volume  $V_\delta \gg M_s^{-\delta}$  where  $M_s$  is the string scale. Our goal is to study how the resulting theory evolves as we take  $V_\delta \rightarrow \infty$ .

For simplicity, we shall start by considering *untwisted* compactifications, temporarily deferring our analysis of situations with twisted compactifications (such as arise in orbifold compactifications) to Sec. III F. We also remark that in the case of four-dimensional closed strings, we would generally be compactifying from ten dimensions to four dimensions. There are therefore a total of six compactified dimensions, and we choose  $\delta$  to represent the number of such dimensions whose characteristic sizes we wish to consider growing increasingly large. Thus  $0 < \delta \leq 6$ . Moreover, in keeping with the observation at the end of the previous section, we shall focus on the cases with  $\delta \in 2\mathbb{Z} > 0$ .

Within such modular-invariant theories, we shall concentrate on physical quantities  $\zeta$  for which the corresponding one-loop contributions  $\langle \mathcal{X} \rangle$  are finite for all  $V_\delta$ . In four spacetime dimensions,  $\langle \mathcal{X} \rangle$  can generally have at most a logarithmic divergence. From direct inspection of Eq. (2.6), and as noted in Eq. (2.19), we see that such a logarithmic divergence is proportional to  $\text{Str}_{M=0} \mathbb{X}_2$  where the supertrace is restricted to the massless  $\mathbb{X}_2$ -charged states in the string spectrum. Thus, until further notice, we shall concentrate on quantities  $\zeta$  for which

$$\text{Str}_{M=0} \mathbb{X}_2 = 0. \quad (3.1)$$

We emphasize that for our purposes this constraint is not exceedingly restrictive, since our primary interest in this paper is in the effects of decompactifications, and as we shall see this question turns out to be largely independent of whether the amplitudes are finite or infinite. Therefore, for simplicity and concreteness, we shall focus on amplitudes which are finite, and hence satisfy the constraint in Eq. (3.1). Likewise, in most phenomenologically interesting string models, it can be difficult to produce a massless  $\mathbb{X}_2$ -charged state, since the same worldsheet excitations that produce a nonzero  $\mathbb{X}_2$  charge will also render the corresponding state massive. We shall therefore assume that Eq. (3.1) holds for the rest of this section. Of course, for physical quantities  $\zeta$  for which this condition is not satisfied, it would be necessary to introduce a regulator function  $\mathcal{G}$ , as described in Sec. II B. This would introduce several subtleties into the proof of our main theorem but will not alter the main result. Such situations will be discussed later in this paper.

Given that we are temporarily focusing on untwisted compactifications, the  $\mathcal{X}$ -inserted partition function  $Z_{\mathcal{X}}^{(4)}$  of our compactified four-dimensional theory can be factorized, i.e.,

$$Z_{\mathcal{X}}^{(4)} = Z_{\mathcal{X}}^{(\text{base})} \cdot Z_{\text{KK/winding}}, \quad (3.2)$$

where  $Z_{\text{KK/winding}}$  is the trace over all of the Kaluza-Klein (KK) and winding modes associated with the  $\delta$  large compactified extra dimensions and where the ‘‘base’’ partition function  $Z_{\mathcal{X}}^{(\text{base})}$  can be written as

$$Z_{\mathcal{X}}^{(\text{base})} = \tau_2^{-1} \sum'_{mn} a_{mn} \mathcal{X}_{mn} \bar{q}^m q^n \quad (3.3)$$

where  $\sum'_{mn}$  indicates a sum over the states *excluding* the KK and winding modes associated with the  $\delta$  large dimensions. In other words, our original theory can be viewed as a ‘‘base’’ theory tensored with a cloud of KK/winding states stemming from the compactification, with each state in the base theory accruing the same set of KK/winding excitations.

In general,  $Z_{\mathcal{X}}^{(\text{base})}$  contains the information regarding our specific theory independently of the compactification. This is thus the portion of the original  $\mathcal{X}$ -inserted partition function that depends on the particular operator insertion  $\mathcal{X}$  but which is independent of the compactification volume  $V_\delta$ . By contrast,  $Z_{\text{KK/winding}}$  contains the information regarding the specific geometry associated with the large compactified dimensions. As such,  $Z_{\text{KK/winding}}$  is independent of  $\mathcal{X}$  but depends on  $V_\delta$ . For example, if we specialize to  $\delta = 1$  and define the dimensionless radius  $\tilde{R} \equiv M_s R = R/\sqrt{\alpha'}$ , we have

$$\begin{aligned} Z_{\text{KK/winding}} &= \sum_{\tilde{m}, \tilde{n}} \bar{q}^{(\tilde{m}/\tilde{R} - \tilde{n})^2/4} q^{(\tilde{m}/\tilde{R} + \tilde{n})^2/4} \\ &= \sum_{\tilde{m}, \tilde{n}} e^{-\pi\tau_2(\tilde{m}^2/\tilde{R}^2 + \tilde{n}^2\tilde{R}^2)} e^{2\pi i\tau_1 \tilde{m} \tilde{n}} \\ &= Z_{\text{circ}}/\sqrt{\tau_2} \end{aligned} \quad (3.4)$$

where  $\tilde{m}$  and  $\tilde{n}$  respectively index the KK (momentum) and winding modes associated with this large extra dimension and where  $Z_{\text{circ}}$  is the modular-invariant circle partition function

$$Z_{\text{circ}}(\tilde{R}, \tau) \equiv \sqrt{\tau_2} \sum_{\tilde{m}, \tilde{n} \in \mathbb{Z}} \bar{q}^{(\tilde{m}/\tilde{R} - \tilde{n})^2/4} q^{(\tilde{m}/\tilde{R} + \tilde{n})^2/4}. \quad (3.5)$$

In this case we thus see from the middle line of Eq. (3.4) that  $Z_{\text{KK/winding}}$  traces over KK/winding states with masses

$$M_{\tilde{m}, \tilde{n}}^2 = \frac{\tilde{m}^2}{R^2} + \tilde{n}^2 M_s^4 R^2, \quad (3.6)$$

as required. Likewise, for a  $\delta$ -dimensional (square) toroidal compactification we may take  $Z_{\text{KK/winding}} = (Z_{\text{circ}}/\sqrt{\tau_2})^\delta$ .

Even though the masses of the KK states in Eq. (3.6) are the same as we would expect for a five-dimensional field theory compactified on a circle, this summation also includes the contributions from winding modes and is thereby modular invariant, so that the full  $\mathcal{X}$ -inserted

partition function  $Z_{\mathcal{X}}^{(4)}$  in Eq. (3.2) is modular invariant. Indeed, even though the individual factors in Eq. (3.2) are not separately modular invariant, we may reshuffle factors of  $\tau_2$  in order to write

$$Z_{\mathcal{X}}^{(4)} = (\tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})}) \cdot (\tau_2^{\delta/2} Z_{\text{KK/winding}}), \quad (3.7)$$

where now each factor is individually modular invariant. For compactifications of the sort we are discussing, a modular-invariant reshuffling such as that in Eq. (3.7) is completely general, independent of the spacetime geometry. Indeed, for a square  $\delta$ -dimensional toroidal compactification, the final factor is nothing but the modular-invariant sum  $(Z_{\text{circ}})^\delta$ .

Let us now ask what happens as  $V_\delta \rightarrow \infty$ . It is once again easiest to focus on the case of square toroidal compactifications as a guide and ask what happens when the radius associated with these  $\delta$  dimensions becomes large, with  $R^{-1} \ll M_s$  or equivalently  $\tilde{R} \gg 1$ . In the  $M_s R \rightarrow \infty$  limit we can disregard all excited winding-mode states with  $\tilde{n} \neq 0$ , as these states become infinitely massive. Likewise, the KK masses become essentially continuous. In this limit we can then evaluate  $Z_{\text{KK/winding}}$ , obtaining

$$\begin{aligned} Z_{\text{KK/winding}} &\approx \sum_{\tilde{m}} e^{-\pi\tau_2 \tilde{m}^2 / \tilde{R}^2} \\ &= \frac{M_s R}{\sqrt{\tau_2}} \sum_{\ell \in \mathbb{Z}} e^{-\pi\ell^2 (M_s R)^2 / \tau_2} \\ &\approx \frac{M_s R}{\sqrt{\tau_2}}. \end{aligned} \quad (3.8)$$

Note that in passing from the first line of Eq. (3.8) we have employed an exact Poisson resummation, while the passage to the third line then follows by taking the  $R \rightarrow \infty$  limit. Of course, we could have obtained the same results by approximating the sum in the first line as an integral, which would have led to the third line directly.

Likewise, for  $\delta$  orthogonal dimensions of radius  $R$ , we obtain

$$Z_{\text{KK/winding}} \approx \frac{M_s^\delta R^\delta}{\tau_2^{\delta/2}} = \frac{\mathcal{M}^\delta V_\delta}{\tau_2^{\delta/2}} \quad (3.9)$$

where  $\mathcal{M} \equiv M_s/(2\pi)$  is the reduced string scale and where  $V_\delta = (2\pi R)^\delta$  is the compactification volume. Once again, we note that the final expression in Eq. (3.9) is completely general, holding independently of the (factorized) compactification geometry.

However, partition functions of compactified string theories in different spacetime dimensions are generally related via

$$Z^{(D+\delta)} \equiv \lim_{V_\delta \rightarrow \infty} \left[ \frac{1}{\mathcal{M}^\delta V_\delta} Z^{(D)} \right] \quad (3.10)$$

where the theory corresponding to  $Z^{(D)}$  has  $\delta$  large compactification radii with compactification volume  $V_\delta$ . Indeed, as  $V_\delta \rightarrow \infty$ , the partition function  $Z^{(D)}$  develops a divergence which scales as the volume  $V_\delta$  of compactification; dividing out by this volume as in Eq. (3.10) then yields the finite higher-dimensional partition function  $Z^{(D+\delta)}$ . Putting the pieces together, we therefore find that

$$Z_{\mathcal{X}}^{(4+\delta)} = \lim_{V_\delta \rightarrow \infty} \left[ \frac{1}{\mathcal{M}^\delta V_\delta} Z_{\mathcal{X}}^{(4)} \right] = \tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})} \quad (3.11)$$

where  $Z_{\mathcal{X}}^{(\text{base})}$  is given in Eq. (3.3). This observation allows us to identify the “base” factor within our four-dimensional theory in terms of the higher-dimensional theory:

$$Z_{\mathcal{X}}^{(\text{base})} = \tau_2^{\delta/2} Z_{\mathcal{X}}^{(4+\delta)}. \quad (3.12)$$

Note that both sides of this relation are indeed  $V_\delta$ -independent.

As we have seen, Eq. (3.10) relates modular-invariant theories in different dimensions. We shall refer to this equation as a “smoothness” constraint because it ensures that the four-dimensional partition function smoothly becomes a  $(4 + \delta)$ -dimensional partition function in the  $V_\delta \rightarrow \infty$  limit. In this context, we note that Eq. (3.10) directly implies a similar smoothness relation for the corresponding one-loop *amplitudes*. Indeed, defining

$$\langle \mathcal{X} \rangle^{(4+\delta)} \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z_{\mathcal{X}}^{(4+\delta)}, \quad (3.13)$$

we have

$$\langle \mathcal{X} \rangle^{(4+\delta)} = \lim_{V_\delta \rightarrow \infty} \frac{1}{\mathcal{M}^\delta V_\delta} \langle \mathcal{X} \rangle^{(4)}. \quad (3.14)$$

However, while Eq. (3.10) implies Eq. (3.14), the converse is not true. Indeed, while two equal partition functions lead to identical amplitudes, identical amplitudes only imply equality of the corresponding partition functions modulo functions  $f(\tau, \bar{\tau})$  whose  $\mathcal{F}$ -integrals vanish. Indeed, many such nonzero functions  $f(\tau, \bar{\tau})$  with vanishing  $\mathcal{F}$ -integrals are known to exist [25–27]. We also note that  $\langle \mathcal{X} \rangle^{(4)}$  scales as the compactification volume  $V_\delta$  as  $V_\delta \rightarrow \infty$ ; dividing out by  $V_\delta$  as in Eq. (3.14) then leads to a finite higher-dimensional amplitude  $\langle \mathcal{X} \rangle^{(4+\delta)}$ . In this connection we note that this divergence of  $\langle \mathcal{X} \rangle^{(4)}$  as  $V_\delta \rightarrow \infty$  is associated with a mere overall multiplicative factor. In particular, it is *not* associated with the modular integration of  $Z^{(4)}$  over the fundamental domain  $\mathcal{F}$  (such as might arise due to certain massless or tachyonic states).

The (re)emergence of a higher-dimensional theory in the large-volume limit is certainly not a surprise. Indeed, geometric decompactification is an intrinsically smooth

and continuous process. *However*, the extra factor  $\tau_2^{-\delta/2}$  which appears in Eq. (3.11) is of critical importance. This extra factor indicates that if  $Z_{\mathcal{X}}^{(\text{base})}$  has a  $\tau_2$ -dependent prefactor  $\tau_2^{-1}$ , as appropriate for a four-dimensional theory, then  $Z_{\mathcal{X}}^{(4+\delta)}$  has a  $\tau_2$ -dependent prefactor  $\tau_2^k$  where  $k = 1 - D/2$  with  $D = 4 + \delta$ . The appearance of the new factor  $\tau_2^{-\delta/2}$  in Eq. (3.11) thus reflects the change in dimensionality of any modular-invariant theory when an extra uncompactified spacetime dimension comes into existence.

It is here that we witness the fundamental clash between the smoothness of the decompactification process and the discrete integer nature of the number of uncompactified spacetime dimensions (or equivalently the half-integer nature of the modular weight  $k \in \mathbb{Z}/2$ ). Indeed, while the  $V_\delta \rightarrow \infty$  limit is essentially a smooth one as far as the resulting physics is concerned, the powers of  $\tau_2$  change in this limit in what is ultimately a discontinuous way according to Eq. (2.2).

It is important to properly understand the nature of this discontinuity. Toward this end, let us revisit Eq. (2.1). We may regard the form of this expression as the “canonical” form for a partition function. Indeed, this form consists of a discrete double power series in  $q$  and  $\bar{q}$ , where  $q \equiv e^{2\pi i\tau}$ , along with an overall factor of  $\tau_2$  raised to a certain power  $k$ . The canonical form of the partition function is of utmost importance because it is only in this form that one can read off a value of  $k$  which can be interpreted as a modular weight—indeed, the same modular weight  $k$  which appears throughout the Rankin-Selberg procedure. In other words, it is only when we cast our partition function into the canonical form that we expose the true modular weight  $k$  of our theory.

Such a partition function can also depend on a compactification radius  $R$ , which is a continuous variable. As we have stated above, the underlying physics of our theory must have a smooth  $R \rightarrow \infty$  limit. Indeed, for every value of  $R$  (including infinity), it is possible to recast our partition function into the canonical form in Eq. (2.1); moreover, for every *finite* value of  $R$ , the value of  $k$  that appears in the canonical form remains the same (equalling  $-1$  for four-dimensional theories). However, in the  $R \rightarrow \infty$  limit, the value of  $k$  that appears in the canonical form jumps to a new value. For example, in the case of a four-dimensional theory with a single extra dimension, we now have  $k = -3/2$  in the  $R \rightarrow \infty$  limit, consistent with Eq. (2.2). This is the “discontinuity” to which we are referring.

We stress that the underlying physics is not discontinuous in this limit; it is merely *the passage to the canonical form comprising a discrete double power series* that becomes discontinuous. Indeed, this discontinuity arises from the fact that in the infinite-radius limit  $Z_{\text{KK/winding}}$  can no longer be written in the same canonical form as for finite radius. Instead, what happens at infinite radius is that our

discrete spectrum becomes continuous. As a result, in this limit,  $Z_{\text{KK/winding}}$  takes the form of a divergent volume factor multiplying a volume-independent expression. However, this volume-independent expression is in the canonical form, but now with a different value of  $k$ . The passage to the higher-dimensional theory as in Eq. (3.10) then eliminates this volume factor, but leaves us with a new canonical form with a new value of  $k$ .

This, then, is the fundamental clash between modular invariance and the process of decompactification. We know that the process of decompactification must ultimately be smooth, even in the decompactification limit. On the other hand, the value of  $k$  within the canonical form changes in a discontinuous way in the decompactification limit—with an extra factor of  $\tau_2^{-\delta/2}$  appearing in Eq. (3.11)—and we know that  $k$  is a quantity which is absolutely fundamental in describing the modular properties of our theory. How then can these two features be reconciled?

Before proceeding further, we note that this is not the first time such clashes have arisen within string theory, or even simply within conformal field theory. For example, let us consider the case of a boson compactified on a circle of radius  $R$ , as discussed in Ref. [28]. If  $R$  is rational, it can be expressed in lowest form as  $p/q$  for some integers  $(p, q)$ , and the resulting decomposition of the partition function into left- and right-moving CFT characters depends critically on the values of  $p$  and  $q$ . Thus, as we sweep through rational values of  $R$ , it would seem that the corresponding partition functions—and therefore the properties of the resulting CFTs—will vary hugely and discontinuously. The existence of irrationals amongst the rationals only introduces further potential discontinuities into the mix. Yet we know that the physics must ultimately be smooth as we vary  $R$ .

How can this clash be resolved? In Ref. [28], it was shown that modular invariance—specifically the relevant GSO projections between the left- and right-moving sectors of the theory—must always connect the different CFTs in a way that is responsible for restoring continuity to all physical amplitudes as a function of  $R$ . In our case, by contrast, we are dealing with a clash between modular invariance (specifically a discontinuous change in the modular weight) and the process of decompactification. What then is the analogous resolution to this puzzle?

### B. Sharpening the clash

In order to address this issue, let us first sharpen this question still further. We shall do this by examining the consequences of this extra factor of  $\tau_2^{-\delta/2}$  in Eq. (3.11). Although there are several ways in which we might incorporate this factor into our analysis, the most straightforward way is to bundle it with the leading prefactor  $\tau_2^{-1}$  in Eq. (3.11). As noted above, this then produces the net prefactor  $\tau_2^{1-D/2}$  that we expect of a fully  $D$ -dimensional theory, with  $D = 4 + \delta$ .

However, what is perhaps less obvious is how this new factor of  $\tau_2^{-\delta/2}$  affects our results for the amplitude  $\langle \mathcal{X} \rangle$ . To see this, let us now proceed to apply the Rankin-Selberg procedure in order to analyze the one-loop amplitudes  $\langle \mathcal{X} \rangle^{(4)}$  and  $\langle \mathcal{X} \rangle^{(4+\delta)}$  that correspond to  $Z_{\mathcal{X}}^{(4)}$  and  $Z_{\mathcal{X}}^{(4+\delta)}$  in Eqs. (3.2) and (3.11) respectively. The corresponding  $g$ -functions  $g(\tau_2) \equiv \int_{-1/2}^{1/2} d\tau_1 Z(\tau, \bar{\tau})$  can then be written as  $g^{(4)} = \tau_2^{-1} S^{(4)}$  and  $g^{(4+\delta)} = \tau_2^{-1-\delta/2} S^{(4+\delta)}$  where

$$\begin{aligned} S^{(4)} &\equiv \int_{-1/2}^{1/2} d\tau_1 \left[ \left( \sum'_{mn} a_{mn} \mathcal{X}_{mn} \bar{q}^m q^n \right) Z_{\text{KK/winding}} \right] \\ S^{(4+\delta)} &\equiv \sum'_n a_{nn} \mathcal{X}_{nn} (\bar{q}q)^n \end{aligned} \quad (3.15)$$

and where we remind the reader that the primes on the summations  $\sum'_{mn}$ , just as in Eq. (3.3), indicate sums over the states *excluding* the KK and winding modes associated with the  $\delta$  large dimensions. Indeed, we shall generally use primes to indicate quantities uniquely associated with the “base” theory in Eq. (3.3) rather than the full theory which also includes the compactification factor  $Z_{\text{KK/winding}}$ .

Following Eq. (2.27), we can then expand  $S^{(4)}$  and  $S^{(4+\delta)}$  in powers of  $\tau_2$  as  $\tau_2 \rightarrow 0$ , i.e.,

$$\begin{aligned} S^{(4)} &\sim \sum_j C_j \tau_2^j \\ S^{(4+\delta)} &\sim \sum_j C'_j \tau_2^j, \end{aligned} \quad (3.16)$$

where the  $C$  and  $C'$  coefficients correspond to  $S^{(4)}$  and  $S^{(4+\delta)}$  respectively. Indeed, given the expression for  $S^{(4+\delta)}$  in Eq. (3.15), we immediately have

$$\begin{aligned} C'_0 &= \text{Str}' \mathcal{X} \\ C'_1 &= \text{Str}' \frac{d\mathcal{X}}{d\tau_2} - \pi\alpha' \text{Str}'(\mathcal{X}M^2) \\ C'_2 &= \frac{1}{2} \text{Str}' \frac{d^2\mathcal{X}}{d\tau_2^2} - \pi\alpha' \text{Str}' \left( \frac{d\mathcal{X}}{d\tau_2} M^2 \right) + \frac{1}{2} \pi^2 (\alpha')^2 \text{Str}'(\mathcal{X}M^4) \\ &\vdots \\ C'_n &= \frac{1}{n!} \text{Str}' D_{\tau_2}^n \mathcal{X} \end{aligned} \quad (3.17)$$

where  $D_{\tau_2}$  is defined in Eq. (2.31) and where the primes on the supertraces in Eq. (3.17) continue to indicate that the KK and winding states associated with the  $\delta$  large dimensions are excluded.

Given the  $C$ - and  $C'$ -coefficients, the Rankin-Selberg procedure outlined in Sec. II C then tells us that

$$\begin{aligned}\langle \mathcal{X} \rangle^{(4)} &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^{-1} S^{(4)} \\ \langle \mathcal{X} \rangle^{(4+\delta)} &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^{-1-\delta/2} S^{(4+\delta)}.\end{aligned}\quad (3.18)$$

The presumed finiteness of  $\langle \mathcal{X} \rangle^{(4)}$  then allows us to conclude that the  $C$ -coefficients satisfy

$$\begin{cases} C_0 = 0 \\ \langle \mathcal{X} \rangle^{(4)} = \frac{\pi}{3} C_1, \end{cases}\quad (3.19)$$

as expected for any four-dimensional theory. Likewise, for the  $C'$ -coefficients, the corresponding finiteness of  $\langle \mathcal{X} \rangle^{(4+\delta)}$  allows us to obtain results which depend critically on  $\delta$ :

$$\begin{aligned}\delta = 2: & \begin{cases} C'_0 = C'_1 = 0 \\ \langle \mathcal{X} \rangle^{(6)} = \frac{\pi}{3} C'_2 \end{cases} \\ \delta = 4: & \begin{cases} C'_0 = C'_1 = C'_2 = 0 \\ \langle \mathcal{X} \rangle^{(8)} = \frac{\pi}{3} C'_3 \end{cases} \\ \delta = 6: & \begin{cases} C'_0 = C'_1 = C'_2 = C'_3 = 0 \\ \langle \mathcal{X} \rangle^{(10)} = \frac{\pi}{3} C'_4. \end{cases}\end{aligned}\quad (3.20)$$

Indeed, these results are all manifestations of the misaligned supersymmetry that governs the spectra of modular-invariant string theories in different dimensions, even without spacetime supersymmetry. Moreover, these results provide a direct illustration of our earlier assertion that the spectra of modular-invariant string theories exhibit increasingly many internal cancellations as the spacetime dimension increases.

These results allow us to rephrase and sharpen the “discontinuity” that occurs for our compactified string theory as  $V_\delta \rightarrow \infty$ . To see this, let us recall, as in Eq. (3.2), that our compactified string theory consists of two components tensored together: a “base” theory and a “cloud” of KK/winding-mode excitations. We have also seen in Eq. (3.12) that the base theory is nothing but the higher-dimensional theory prior to compactification. Finally, we have been assuming that the overall physical amplitude  $\langle \mathcal{X} \rangle$  associated with our theory remains finite for all  $V_\delta$ . Given these assumptions, we can ask what constraints must be satisfied by our theory as a function of  $V_\delta$ . In general, there are two classes of constraints:

- (i)  $C$ -constraints that govern the entire spectrum of the four-dimensional string model; and
- (ii)  $C'$ -constraints that govern that portion of the spectrum associated with the “base” theory alone.

For any finite  $V_\delta$ , the finiteness of our overall amplitude  $\langle \mathcal{X} \rangle$  (i.e., the finiteness of  $\langle \mathcal{X} \rangle^{(4)}$ ) requires—at a bare minimum—that

$$\begin{cases} \bullet C_0 = 0 \\ \bullet C'_j \text{ arbitrary.} \end{cases}\quad (3.21)$$

However, let us now consider what happens as  $V_\delta \rightarrow \infty$ . In this limit, the overall amplitude  $\langle \mathcal{X} \rangle^{(4)}$  technically accrues a “spurious” divergence due to the infinite volume factor in Eq. (3.14). However, as  $V_\delta \rightarrow \infty$ , our theory is now in higher dimensions. This means, according to Eq. (3.14), that we should divide out by this volume in order to continue to obtain the corresponding amplitude  $\langle \mathcal{X} \rangle$ . Indeed, the resulting amplitude is now nothing but  $\langle \mathcal{X} \rangle^{(4+\delta)}$ . Thus, as  $V_\delta \rightarrow \infty$ , the continued finiteness of our overall amplitude  $\langle \mathcal{X} \rangle$  now translates to the finiteness of  $\langle \mathcal{X} \rangle^{(4+\delta)}$ , which in turn requires

$$\{\bullet C'_0 = C'_1 = \dots = C'_{\delta/2} = 0.\quad (3.22)$$

This sudden shift in the constraints on our string model as  $V_\delta \rightarrow \infty$  is the manifestation of the discontinuity we have been discussing all along, namely the clash between the decompactification limit and the requirements of modular invariance.

It is easy to understand how and why these different sets of constraints arise. In the discussion above, we have let  $\langle \mathcal{X} \rangle$  represent the physical amplitude of our theory. When  $V_\delta$  is finite, this quantity is nothing other than  $\langle \mathcal{X} \rangle^{(4)}$ . However, when  $V_\delta$  is infinite, this quantity is nothing other than  $\langle \mathcal{X} \rangle^{(4+\delta)}$ . What we are demanding is simply that this transition as  $V_\delta \rightarrow \infty$  be a *smooth* one, with no discontinuity in the physical amplitude  $\langle \mathcal{X} \rangle$ . If  $\langle \mathcal{X} \rangle$  is finite for *all*  $V_\delta$  [where we have already compensated for the spurious volume divergence via Eq. (3.14)], then we are saying that our theory has no choice but to satisfy the constraints in Eq. (3.21) for all finite  $V_\delta$  and to satisfy the constraints in Eq. (3.22) for infinite  $V_\delta$ . This sudden shift in the constraints on our string model as  $V_\delta \rightarrow \infty$  is the manifestation of the apparent discontinuity we are seeking to resolve.

### C. Resolving the clash: Our fundamental theorem

Ultimately, there is only one way in which the constraints in Eq. (3.21) for finite  $V_\delta$  can be smoothly reconciled with those in Eq. (3.22) for infinite  $V_\delta$ : our string theory must actually satisfy the more stringent constraints

$$\begin{cases} \bullet C_0 = 0 \\ \bullet C'_0 = C'_1 = \dots = C'_{\delta/2} = 0 \end{cases}\quad (3.23)$$

for *all* compactification volumes  $V_\delta$ , finite or infinite. Indeed, this is tantamount to demanding that the extra constraints  $C'_j = 0$  for all  $0 \leq j \leq \delta/2$  apply not only in the  $V_\delta \rightarrow \infty$  limit, but rather *for all*  $V_\delta$ . Note that in making this demand we are not *introducing* a new set of constraints for string models with decompactification limits. Instead, we are simply asserting that such string models must *already* have been satisfying these constraints, even if these constraints had not been explicitly noticed before.

Indeed, *it is these properties that allow the decompactification limits to exist.*

This assertion represents the content of our theorem. Specifically, we have the following:

*Theorem.* Any four-dimensional closed string theory which can be realized as a geometric compactification from a higher-dimensional string theory will inherit the precise stricter internal cancellations of the higher-dimensional theory despite the compactification.

We shall prove this theorem in Sec. III D. In this connection, we remind the reader that we have been limiting our discussion here to theories whose partition functions can be factored as in Eq. (3.2)—i.e., theories whose compactifications are untwisted. However, as we shall soon discuss, the above theorem can actually be trivially generalized to apply to *any* compactification, twisted or untwisted.

In Sec. IV, we shall see why we may regard this as a nonrenormalization theorem. For now, however, we simply note that this theorem may also conversely be viewed as providing an important constraint on the construction of compactified string models. Indeed, as already noted, our compactified string theory consists of a “base” theory tensored with a cloud of KK/winding states. We might then ask whether we can tensor such a cloud of KK/winding states to *any* base theory. In a field-theoretic context, the answer is yes. However, in string theory, the requirements of modular invariance imply that we cannot do this unless certain (primed) supertraces in the base theory vanish. These are the primed supertraces associated with the  $C'_j$ -coefficients for  $0 \leq j \leq \delta/2$ . Indeed this is the only way in which we can smoothly and self-consistently absorb the extra powers of  $\tau_2$  which arise in the decompactification limit.

We conclude this discussion of our theorem with one final comment. In general, while the  $C_j$ -coefficients are messy functions of compactification radius and geometry, the  $C'_j$ -coefficients are by definition independent of any details of compactification. For example, a six-dimensional theory compactified on a two-torus and the same six-dimensional theory compactified on a two-sphere will give rise to *distinct* four-dimensional theories. However, these four-dimensional theories will nevertheless share the same  $C'_j$ -coefficients because they flow to the same six-dimensional theory as the volume  $V_\delta$  becomes large. Thus, the space of compactified four-dimensional theories can be separated into different *equivalence classes* based on their internal  $C'$ -constraints—i.e., equivalence classes which depend on the higher-dimensional theories to which they flow at large volume.

#### D. Proving the theorem

To prove our theorem, we shall begin by proving that any four-dimensional closed string theory which can be realized as a geometric compactification from a  $(4 + \delta)$ -dimensional

string theory with arbitrary compactification volume  $V_\delta$  satisfies the constraints given in Eq. (3.23) rather than those given in Eq. (3.21). To do this, let us study the relationship between the  $C_j$ -coefficients and the  $C'_j$ -coefficients.

Given that these coefficients respectively describe our original four-dimensional theory and the “base” of that theory, and given that this base is nothing but the original  $(4 + \delta)$ -dimensional theory, any relationship between these two sets of coefficients must stem from a relationship between the compactified and uncompactified theories. However, we have already seen such a relationship: this is our “smoothness” constraint in Eqs. (3.14). Performing a  $\tau_1$ -integration of both sides of this relation over  $-1/2 < \tau_1 \leq 1/2$ , inserting the expansions in Eq. (3.16), and matching terms with equal powers of  $\tau_2$  then yields the relation

$$C'_j = \lim_{V_\delta \rightarrow \infty} \frac{1}{\mathcal{M}^\delta V_\delta} C_{j-\delta/2} \quad \text{for all } j. \quad (3.24)$$

As discussed above, the shifting of the  $j$ -index reflects the extra powers of  $\tau_2$  that emerge upon the decompactification of the large spacetime dimensions. Indeed, this index shifting is required by modular invariance and our smoothness requirement.

We have seen that our smoothness constraint on the partition functions in Eq. (3.10) leads directly to the smoothness constraint in Eq. (3.14) on the corresponding *amplitudes*  $\langle \mathcal{X} \rangle^{(4)}$  and  $\langle \mathcal{X} \rangle^{(4+\delta)}$ . Indeed, it is the finiteness of  $\langle \mathcal{X} \rangle^{(4)}$  that implies the auxiliary condition that  $C_j = 0$  for all  $j < 1$ , which includes the constraint  $C_0 = 0$ . From the relation in Eq. (3.24) we then find

$$C'_j = 0 \quad \text{for all } j < 1 + \delta/2. \quad (3.25)$$

This result is consistent with the results quoted in Eq. (3.20), and is tantamount to the assertion that if  $\langle \mathcal{X} \rangle^{(4)}$  is finite (after dividing out, of course, the overall volume factor which diverges in the  $V_\delta \rightarrow \infty$  limit), then  $\langle \mathcal{X} \rangle^{(4+\delta)}$  is also finite. In other words, the  $\mathcal{X}$ -amplitude of our four-dimensional compactified theory cannot suddenly grow a new divergence in the  $V_\delta \rightarrow \infty$  limit. This then completes the proof of our theorem.

In fact, we can even push things one step further. Thus far, we have seen that we have two kinds of constraints: our  $C$ -constraints which govern the entire four-dimensional theory and our  $C'$ -constraints which govern the “base” portion of that theory (or equivalently which govern its higher-dimensional decompactification limit). However, we shall now demonstrate that there is in fact a universal relation between these two groups of constraints. Indeed, this relation will apply to any theory that has a decompactification limit regardless of the degree to which its partition function factorizes.

To proceed let us begin with two fundamental observations:

- (i) The result in Eq. (3.24) does not depend on the compactification geometry. All that is assumed is that the partition function of any string theory in  $D$  spacetime dimensions has a leading prefactor of  $\tau_2^k$  where  $k = 1 - D/2 = -1 - \delta/2$  where  $\delta = D - 4$ . This is a general result for any compactification. This also does not assume an *untwisted* compactification (i.e., it does not assume that the four-dimensional partition function  $Z^{(4)}$  factorizes), for the same reason. Indeed, the  $C$ -expansion in Eq. (3.16) is completely general regardless of the precise form of the quantity  $S^{(4)}$  in Eq. (3.15) as long as  $S^{(4)}$  corresponds to a four-dimensional theory, so that  $g^{(4)}(\tau_2) = \tau_2^{-1} S^{(4)}$ .
- (ii) The results in Eqs. (3.19) and (3.20) are also completely general, following from the same feature as described above.

Given these observations, our task is now to relate the  $C_j$ -constraints from the  $C'_j$ -constraints—not just at infinite volume but even at finite volume.

The easiest way to proceed is to consider the *difference* between our partition functions

$$\Delta Z_{\mathcal{X}} \equiv Z_{\mathcal{X}}^{(4+\delta)} - \frac{1}{\mathcal{M}^\delta V_\delta} Z_{\mathcal{X}}^{(4)}. \quad (3.26)$$

Note that in constructing this difference we are *not* taking the  $V_\delta \rightarrow \infty$  limit; thus this difference is a function of  $V_\delta$ . By considering only the difference in this way we avoid making any assumptions about the behavior of  $\langle \mathcal{X} \rangle^{(4)}$  at finite  $V_\delta$ . In this connection we note that the difference of two partition functions is not necessarily itself the partition function of any self-consistent string model (see, e.g., Refs. [26,27]). However, such a property is not required for our proof.

Given this definition for  $\Delta Z_{\mathcal{X}}$ , we can then define the corresponding amplitude

$$\langle \Delta \mathcal{X} \rangle = \int \frac{d^2 \tau}{\tau_2} \Delta Z_{\mathcal{X}}, \quad (3.27)$$

the corresponding  $g$ -function

$$g_{\Delta \mathcal{X}} \equiv \int d\tau_1 Z_{\Delta \mathcal{X}}, \quad (3.28)$$

and the corresponding sum

$$S_{\Delta \mathcal{X}} \equiv \tau_2 g_{\Delta \mathcal{X}}. \quad (3.29)$$

We can also consider expanding  $S_{\Delta \mathcal{X}}$  in powers of  $\tau_2$  in the  $\tau_2 \rightarrow 0$  limit, i.e.,

$$S_{\Delta \mathcal{X}} = \tau_2^{-\delta/2} \sum_j C''_j \tau_2^j \quad \text{as } \tau_2 \rightarrow 0, \quad (3.30)$$

thereby defining a new set of  $C''_j$ -coefficients.

Let us now discuss the finiteness of  $\Delta Z$ . Of course, we learn from Eq. (3.10) that  $\Delta Z_{\mathcal{X}} \rightarrow 0$  as  $V_\delta \rightarrow \infty$ . However, in order for this limit to exist, we also learn that  $\Delta Z_{\mathcal{X}}$  must be *finite* for large  $V_\delta$  (i.e., for  $\mathcal{M}^\delta V_\delta \gg 1$ ). Thus, we see that  $g_{\Delta \mathcal{X}} \rightarrow 0$  as  $V_\delta \rightarrow \infty$  and that  $g_{\Delta \mathcal{X}}$  remains finite for  $\mathcal{M}^\delta V_\delta \gg 1$ . Indeed, these latter assertions follow because the  $\tau_1$ -integration in Eq. (3.28) is incapable of producing a new divergence, given that this integration merely selects the zero mode of the partition-function Fourier series. Likewise, we find that  $S_{\Delta \mathcal{X}} \rightarrow 0$  as  $V_\delta \rightarrow \infty$  and that  $S_{\Delta \mathcal{X}}$  remains finite for  $\mathcal{M}^\delta V_\delta \gg 1$ .

Because these quantities are all finite, the expression for  $\Delta Z$  in terms of the difference between  $Z^{(4+\delta)}$  and  $Z^{(4)}$  allows us to write  $g_{\Delta \mathcal{X}}$  and  $S_{\Delta \mathcal{X}}$  as analogous differences, and thereby ultimately express  $C''_j$  in terms of  $C_j$  and  $C'_j$ . Following this chain of steps, we thus have

$$C''_j = C'_j - \frac{1}{\mathcal{M}^\delta V_\delta} C_{j-\delta/2}. \quad (3.31)$$

Once again we stress that the  $C''_j$  coefficients are generally functions of  $V_\delta$ , with these relations holding for all  $\mathcal{M}^\delta V_\delta \gg 1$ . Likewise, these relations hold as functions of  $\tau_2$  for all  $\tau_2$ .

The final step of our analysis rests on the properties of  $\langle \Delta \mathcal{X} \rangle$ . As discussed below Eq. (3.22), the smoothness of the  $V_\delta \rightarrow \infty$  limit requires that  $\langle \Delta \mathcal{X} \rangle$  be finite for large  $V_\delta$ . Thus, we can take the Rankin-Selberg transform of the finite amplitude  $\langle \Delta \mathcal{X} \rangle$ , i.e.,

$$\langle \Delta \mathcal{X} \rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} [\tau_2^{-1} S_{\Delta \mathcal{X}}(\tau_2)] \quad (3.32)$$

to find that

$$C''_j = 0 \quad \text{for all } j < 1 + \delta/2. \quad (3.33)$$

Indeed, it is the fact that our relations hold for all  $\tau_2$  which enables us to take the  $\tau_2 \rightarrow 0$  limit without difficulty. It then follows that

$$C'_j = \frac{1}{\mathcal{M}^\delta V_\delta} C_{j-\delta/2} \quad \text{for all } j < 1 + \delta/2. \quad (3.34)$$

This is the result we have been seeking. It provides a direct relationship between the  $C_j$  and  $C'_j$  coefficients for  $j < 1 + \delta/2$  and thereby relates our different sets of constraints to each other. It is important to note that Eq. (3.34) is different from Eq. (3.24) because it holds *regardless* of the compactification volume  $V_\delta$ . On the other hand, it holds only for  $j < 1 + \delta/2$ .

Given the result in Eq. (3.34), we see that

$$C'_{\delta/2} = 0 \Rightarrow C_0 = 0, \quad (3.35)$$

with  $C_j$  for all  $j < 0$  vanishing as well. This also implies that

$$\langle \Delta \mathcal{X} \rangle = \langle \mathcal{X} \rangle^{(4+\delta)} - \frac{1}{\mathcal{M}^\delta V_\delta} \langle \mathcal{X} \rangle^{(4)}, \quad (3.36)$$

Indeed, although the left side of this equation is finite, any divergences that appear within the expressions on the right side must be identical so that they cancel in the difference.

Note that Eq. (3.35) holds for all volumes  $V_\delta$ . Indeed, there is only one possible exception to this conclusion. In particular, as  $V_\delta$  becomes smaller, it is possible that a physical, on-shell tachyon might appear. However, the appearance of such a tachyon would signify a breakdown of our compactified theory and automatically result in divergent one-loop amplitudes in any case. This would therefore correspond to taking our theory to a point at which it becomes ill-defined. Thus, we conclude that these results hold for all volumes  $V_\delta$  which correspond to tachyon-free compactifications.

These results provide an additional perspective on our theorem. As we have seen, our theorem states that any four-dimensional string model with a *bona fide* decompactification limit satisfies not only a  $C$ -constraint  $C_0 = 0$  but also a set of additional new  $C'$ -constraints. However, we now see that this  $C$ -constraint can be *replaced* by the  $C'$ -constraints without any loss of generality. Thus the  $C'$  constraints are not only *necessary* (as implied by our theorem) but also *sufficient*. Indeed, any model which satisfies our new  $C'$ -constraints will already satisfy our  $C_0 = 0$  constraint.

Of course, the  $C'$ -constraints that we have discovered here go beyond the  $C$ -constraint that was already known in Ref. [2]. Indeed, as originally discussed in Ref. [2], all four-dimensional closed string theories must satisfy the  $C_0 = 0$  constraint of Eq. (3.19) as long as they are finite (i.e., free of on-shell physical tachyons). However, what we are now learning from our theorem is that if we additionally demand that our four-dimensional theory also have a self-consistent decompactification limit, then this theory must additionally satisfy the  $C'$ -constraints which not only are more powerful than the original  $C_0 = 0$  constraint but even subsume it.

### E. $T$ -volume scaling rule

We shall now present another result which we call the  *$T$ -volume scaling rule*. This result follows from our previous results but now focuses on the first *nonzero* coefficients  $C_1$  and  $C'_{1+\delta/2}$ .

From Eq. (3.36) we find that our original four-dimensional amplitude  $\langle \mathcal{X} \rangle^{(4)}$  is given by

$$\langle \mathcal{X} \rangle^{(4)} = \mathcal{M}^\delta V_\delta [\langle \mathcal{X} \rangle^{(4+\delta)} - \langle \Delta \mathcal{X} \rangle]. \quad (3.37)$$

In principle, this represents a complicated volume dependence for  $\langle \mathcal{X} \rangle^{(4)}$  because  $\langle \Delta \mathcal{X} \rangle$  is itself  $V_\delta$ -dependent even though  $\langle \mathcal{X} \rangle^{(4+\delta)}$  is not. However we know that  $\langle \Delta \mathcal{X} \rangle \rightarrow 0$  at large volume. We therefore expect that

$$\langle \mathcal{X} \rangle^{(4)} \approx \mathcal{M}^\delta V_\delta \langle \mathcal{X} \rangle^{(4+\delta)} \quad \text{for } \mathcal{M}^\delta V_\delta \gg 1. \quad (3.38)$$

In other words, for large volumes, we expect that our amplitude  $\langle \mathcal{X} \rangle^{(4)}$  scales as the volume itself. Indeed, this is nothing but the volume divergence discussed earlier. We also see that the coefficient of this scaling is given by the full amplitude of the original higher-dimensional theory.

As expressed above, however, this result is not consistent with  $T$ -duality. Indeed, from  $T$ -duality considerations we know that our four-dimensional amplitude should scale as the compactification volume not only for very *large* compactification volumes but also for very *small* ones. Toward this end, we seek to define a new kind of (dimensionless) volume—a so-called  *$T$ -volume*  $\tilde{V}_T$ —which is consistent not only with modular invariance but also with  $T$ -duality. One natural proposal for such a quantity would be

$$\tilde{V}_T \stackrel{?}{\equiv} \frac{3}{\pi} \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} \tau_2^{\delta/2} Z_{\text{KK/winding}}. \quad (3.39)$$

This definition has the advantage that it results from a modular-invariant integral and also depends directly on the  $T$ -duality-invariant  $Z_{\text{KK/winding}}$  partition-function factor. Indeed, if we were to naïvely apply the Rankin-Selberg procedure to this integral, we would find

$$\tilde{V}_T \equiv \frac{3}{\pi} \lim_{\tau_2 \rightarrow 0} g_{\text{KK/winding}} \quad (3.40)$$

where

$$g_{\text{KK/winding}} \equiv \int_{-1/2}^{1/2} d\tau_1 \tau_2^{\delta/2} Z_{\text{KK/winding}}. \quad (3.41)$$

However, we can immediately see that the expression in Eq. (3.39) is actually *divergent* for all  $\delta \geq 2$ . Given that  $Z_{\text{KK/winding}}$  has a  $(q, \bar{q})$  expansion that necessarily begins with a nonzero constant term, this divergence arises in the  $\tau_2 \rightarrow \infty$  region due to the extra power  $\tau_2^{\delta/2}$  that was needed for modular invariance. This divergence invalidates the Rankin-Selberg procedure that leads to Eq. (3.40). Indeed, this failure of the Rankin-Selberg procedure can be seen from the fact that Eq. (3.39) is finite only for  $\delta < 2$  while Eq. (3.40) is finite for all  $\delta$ .

Given that Eq. (3.40) is finite for all  $\delta$ , we shall therefore *define*  $\tilde{V}_T$  to be given by Eq. (3.40) rather than by

Eq. (3.39). As we shall see below, this ensures a finite value of  $\tilde{V}_T$  for all  $\delta$ . Moreover, we shall find that it is this definition that leads to meaningful results, and indeed this is all we shall ever need.

This definition for  $\tilde{V}_T$  in Eq. (3.40) provides us with a dimensionless compactification volume which respects  $T$ -duality for the factorized compactifications which have been our focus thus far. The quantity  $\tilde{V}_T$  thereby substitutes for the quantity  $\mathcal{M}^\delta V_\delta$  that we have been writing until now. Furthermore, the overall normalization factor  $3/\pi$  in Eqs. (3.39) and (3.40) ensures that  $\tilde{V}_T = 1$  for the trivial  $\delta = 0$  case in which  $\tau_2^{\delta/2} Z_{\text{KK/winding}} \rightarrow 1$ . Of course, in the special case with  $\delta = 1$ , we find that  $\tilde{V}_T$  is also given by Eq. (3.39). Indeed, for the simple case of compactification on a circle of (dimensionless) radius  $\tilde{R} \equiv M_s R = R/\sqrt{\alpha'}$ , as in Eq. (3.4), we find

$$\tilde{V}_T = \tilde{R} + \frac{1}{\tilde{R}}. \quad (3.42)$$

We thus see that  $\tilde{V}_T \rightarrow \infty$  both at large radius and at small radius, and thereby subsumes both cases in a  $T$ -duality-invariant manner.

Proceeding with this definition of  $\tilde{V}_T$ , we will now show that the coefficient  $C_1$  is indeed given in terms of  $\tilde{V}_T$  by

$$C_1 \approx \tilde{V}_T C'_{1+\delta/2} \quad \text{for } \tilde{V}_T \gg 1, \quad (3.43)$$

or equivalently

$$\langle \mathcal{X} \rangle^{(4)} \approx \tilde{V}_T \langle \mathcal{X} \rangle^{(4+\delta)} \quad \text{for } \tilde{V}_T \gg 1. \quad (3.44)$$

We thus have the following:

*T-volume scaling rule.* Within any four-dimensional modular-invariant closed string theory which can be realized as a geometric compactification from a higher-dimensional string theory, the one-loop amplitude  $\langle \mathcal{X} \rangle^{(4)}$  in the limit of large compactification volume is given by the product of the dimensionless  $T$ -volume of compactification and the corresponding amplitude of the original higher-dimensional theory.

While our proof of this result will hold for un-twisted compactifications, we shall see that it can be easily generalized in order to hold for twisted compactifications as well.

To prove this result, let us recall from Eq. (3.18) that  $\langle \mathcal{X} \rangle^{(4)}$  is given as

$$\langle \mathcal{X} \rangle^{(4)} = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g^{(4)}(\tau_2) \quad (3.45)$$

where

$$g^{(4)}(\tau_2) = \int_{-1/2}^{1/2} d\tau_1 Z_{\mathcal{X}}^{(\text{base})} Z_{\text{KK/winding}}. \quad (3.46)$$

In general, both  $Z_{\mathcal{X}}^{(\text{base})}$  and  $Z_{\text{KK/winding}}$  are double power series in  $q$  and  $\bar{q}$ . Indeed, the latter power series depends on the particular compactification geometry, with an example given in Eq. (3.4) for the case of a one-dimensional compactification on a circle. It is for this reason that  $\langle \mathcal{X} \rangle^{(4)}$  generally depends in a highly nontrivial way on the compactification geometry. Indeed, according to Eq. (3.46) we would need to multiply these two power series together, thereby producing a new power series for the product, whereupon the  $\tau_1$ -integration would project us down to terms with equal coefficients of  $q$  and  $\bar{q}$  in the product. However, because of the fact that our integrand is a *product* of two independent power series, the terms that have equal powers of  $q$  and  $\bar{q}$  in the product need not themselves have equal powers of  $q$  and  $\bar{q}$  for each factor individually. Phrased somewhat differently, if we follow Eq. (3.7) and define

$$g_{\text{base}} \equiv \int_{-1/2}^{1/2} d\tau_1 \tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})} \quad (3.47)$$

along with the definition of  $g_{\text{KK/winding}}$  in Eq. (3.41), we see that

$$g^{(4)} \neq g_{\text{base}} \cdot g_{\text{KK/winding}}. \quad (3.48)$$

Indeed,  $g_{\text{base}}$  and  $g_{\text{KK/winding}}$  are nontrivially *entwined* in forming  $g^{(4)}$ . This phenomenon was discussed in detail in Ref. [17].

To proceed, let us therefore write

$$g^{(4)} = g_{\text{base}} \cdot g_{\text{KK/winding}} + g_{\text{entwined}} \quad (3.49)$$

where  $g_{\text{entwined}}$  represents the “error” term that prevents us from performing a full factorization of  $g^{(4)}$ . Recalling the definition of  $\tilde{V}_T$  in Eq. (3.40), we then find from Eq. (3.45) that

$$\begin{aligned} \langle \mathcal{X} \rangle^{(4)} &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} [g_{\text{base}} \cdot g_{\text{KK/winding}} + g_{\text{entwined}}] \\ &= \frac{\pi}{3} \tilde{V}_T \lim_{\tau_2 \rightarrow 0} g_{\text{base}} + \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g_{\text{entwined}} \\ &= \tilde{V}_T \langle \mathcal{X} \rangle^{(4+\delta)} + \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g_{\text{entwined}}, \end{aligned} \quad (3.50)$$

where we have used Eqs. (3.12) and (3.13) in passing to the final line. Moreover, as promised earlier, we see from the final line of Eq. (3.50) that  $\tilde{V}_T$ —as defined in Eq. (3.40)—is indeed finite because it serves as the proportionality constant between the finite quantities  $\langle \mathcal{X} \rangle^{(4)}$  and  $\langle \mathcal{X} \rangle^{(4+\delta)}$ . Comparison with Eq. (3.37) and replacing  $\mathcal{M}^\delta V_\delta \rightarrow \tilde{V}_T$  then allows us to identify

$$\langle \Delta \mathcal{X} \rangle = -\frac{\pi}{3} \frac{1}{\tilde{V}_T} \lim_{\tau_2 \rightarrow 0} g_{\text{entwined}}. \quad (3.51)$$

Thus we see that  $\langle \Delta \mathcal{X} \rangle$  encapsulates the entwinement between  $g_{\text{base}}$  and  $g_{\text{KK/winding}}$  in the contribution to  $\langle \mathcal{X} \rangle^{(4)}$ . Indeed, contributions from such entwined terms are generally exponentially suppressed relative to those that are unentwined.

We have therefore proven the  $T$ -volume scaling rule, as expressed in Eq. (3.44), with “error” terms that become increasingly small (indeed, exponentially suppressed) as  $\tilde{V}_T \rightarrow \infty$ .

### F. General applicability: Twisted compactifications and multiple constraints

As we have repeatedly stressed, our theorem in Sec. III C has been derived within the context of factorized theories [i.e., theories with factorized partition functions, as in Eq. (3.2)] for which one factor  $Z_{\text{KK/winding}}$  completely describes the compactification geometry. This generally corresponds to unentwined compactifications.

However, there also exist *twisted* compactifications for which this sort of factorization does not arise. These include compactifications on orbifolds; coordinate-dependent Scherk-Schwarz compactifications of the kind discussed in Ref. [29]; and also compactifications involving Wilson-line breaking of gauge symmetries. Likewise, there exist theories (such as Type I strings, or nonperturbative closed strings involving  $D$ -branes) which have some sectors which are modular invariant as well as other sectors which are not modular invariant. It therefore remains to determine the extent to which our theorem applies to such theories as well.

As we shall demonstrate, our theorem does apply to such theories. In particular, our theorem will apply to any modular-invariant portion of any four-dimensional theory which itself becomes a  $(4 + \delta)$ -dimensional theory as a corresponding compactification modulus becomes large.

The issue as to whether or not the partition function of the compactified theory factorizes is not merely an algebraic distinction. Instead, it reflects the manner in which compactification deforms the theory. For an unentwined compactification, the partition function factorizes because the precise KK/winding spectra are the same for each state in the underlying “base” theory. These KK/winding spectra are thus independent of the quantum numbers associated with the states in the base theory. However, for a twisted theory this is no longer true: while one portion of the original base theory is tensored with one set of KK and winding states, another portion of the base theory is tensored with a different set of KK and winding states. In other words, the compactification comes with a “twist” that *correlates* the quantum numbers of the base theory with the quantum numbers associated with the compactification. It is this feature that breaks the factorizability of the partition functions of such theories.

As a result of these observations, it follows that the algebraic structure of the partition function of the resulting

compactified theory depends critically on the numbers and types of twists involved in the compactification. Indeed, one generally obtains a partition function which can be written schematically as the sum of contributions from different *sectors*, i.e.,

$$Z_{\mathcal{X}}^{(4)} = \sum_{\text{sectors } s} Z_s^{(\text{base})} \cdot Z_{\text{KK/winding}}^{(s)} \quad (3.52)$$

where each sector  $s$  is associated with its own  $\mathcal{X}$ -dependent “base function”  $Z_s^{(\text{base})}$  and its own set of KK/winding states associated with  $Z_{\text{KK/winding}}^{(s)}$ .

In order to understand how our theorem can apply in such situations, it will prove simplest to analyze a particular example. Accordingly, for simplicity, we shall consider the case in which our four-dimensional theory is realized as a one-dimensional compactification of a five-dimensional theory, taking our compactification geometry to be that of a circle modded out by a single  $\mathbb{Z}_2$  twist. In this case, we find that the resulting four-dimensional theory has a partition function of the form in Eq. (3.52) with only four sectors, i.e.,  $s = 1, \dots, 4$ .

For this scenario, it is not difficult to identify the resulting  $Z_s^{(\text{base})}$  and  $Z_{\text{KK/winding}}^{(s)}$  factors. Following Ref. [30] while adopting the conventions in Ref. [29], we may take the  $Z_{\text{KK/winding}}^{(s)}$  functions to be none other than  $\mathcal{E}_{0,1/2}/\sqrt{\tau_2}$  and  $\mathcal{O}_{0,1/2}/\sqrt{\tau_2}$ , where these functions are defined to be the same as  $Z_{\text{circ}}$  in Eq. (3.5) except that their summation variables are restricted and shifted as follows:

$$\begin{aligned} \mathcal{E}_0 &= \{\tilde{m} \in \mathbb{Z}, \tilde{n} \text{ even}\} \\ \mathcal{E}_{1/2} &= \left\{ \tilde{m} \in \mathbb{Z} + \frac{1}{2}, \tilde{n} \text{ even} \right\} \\ \mathcal{O}_0 &= \{\tilde{m} \in \mathbb{Z}, \tilde{n} \text{ odd}\} \\ \mathcal{O}_{1/2} &= \left\{ \tilde{m} \in \mathbb{Z} + \frac{1}{2}, \tilde{n} \text{ odd} \right\}. \end{aligned} \quad (3.53)$$

The half-integer modings for  $\tilde{m}$  and the even/odd sensitivity for  $\tilde{n}$  are both related to the  $\mathbb{Z}_2$  nature of the orbifold twist.

Likewise, the corresponding base functions  $Z_s^{(\text{base})}$  in each sector consist of those parts of the original base theory which are either even (unentwined) or odd (twisted) under the action of the orbifold. Specifically, using standard notation, we may identify  $Z_+^+$  as the partition function of the original base theory prior to compactification,  $Z_+^-$  as that of its projection sector,  $Z_-^+$  as that of the corresponding twisted sector, and  $Z_-^-$  as that of the projection sector of the twisted sector. Note that according to the standard conventions for such orbifold-sector partition functions  $Z_{\pm}^{\pm}$  in four dimensions (see, e.g., Ref. [29]) such factors already include factors of  $\tau_2^{-3/2}$ . We can then identify

$$\begin{aligned}
Z_1^{(\text{base})} &= \frac{1}{2} \sqrt{\tau_2} (Z_+^+ + Z_+^-) \\
Z_2^{(\text{base})} &= \frac{1}{2} \sqrt{\tau_2} (Z_-^+ + Z_-^-) \\
Z_3^{(\text{base})} &= \frac{1}{2} \sqrt{\tau_2} (Z_+^+ - Z_+^-) \\
Z_4^{(\text{base})} &= \frac{1}{2} \sqrt{\tau_2} (Z_-^+ - Z_-^-). \quad (3.54)
\end{aligned}$$

Given these identifications, our final orbifolded theory then has a modular-invariant partition function of the form

$$\begin{aligned}
Z_{\mathcal{X}}^{(4)} &= \frac{1}{\sqrt{\tau_2}} \left[ Z_1^{(\text{base})} \mathcal{E}_0 + Z_2^{(\text{base})} \mathcal{E}_{1/2} \right. \\
&\quad \left. + Z_3^{(\text{base})} \mathcal{O}_0 + Z_4^{(\text{base})} \mathcal{O}_{1/2} \right] \quad (3.55)
\end{aligned}$$

where the  $Z_{1,\dots,4}^{(\text{base})}$  functions continue to have the  $\mathcal{X}$ -insertions. In writing Eq. (3.55) we recall that the  $\mathcal{E}, \mathcal{O}$  functions have leading  $\sqrt{\tau_2}$  factors while the  $Z_i^{(\text{base})}$  functions have leading factors  $\tau_2^{-1}$  in four dimensions. Our final result for  $Z_{\mathcal{X}}^{(4)}$  thus has a leading  $\tau_2^{-1}$  factor, as expected.

What will be important for us are the limits of these geometric functions  $\mathcal{E}_{0,1/2}$  and  $\mathcal{O}_{0,1/2}$  as their radii  $\tilde{R} \equiv M_s R = R/\sqrt{\alpha'}$  are taken to be extremely large or small. These can be determined by explicit calculation, yielding

$$\begin{aligned}
\tilde{R} \rightarrow \infty: \quad & \mathcal{E}_0, \mathcal{E}_{1/2} \rightarrow \tilde{R}, \quad \mathcal{O}_0, \mathcal{O}_{1/2} \rightarrow 0 \\
\tilde{R} \rightarrow 0: \quad & \mathcal{E}_0, \mathcal{O}_0 \rightarrow \frac{1}{2\tilde{R}}, \quad \mathcal{E}_{1/2}, \mathcal{O}_{1/2} \rightarrow 0. \quad (3.56)
\end{aligned}$$

From Eq. (3.55) it therefore follows that

$$\begin{aligned}
\tilde{R} \rightarrow \infty: \quad & Z_{\mathcal{X}}^{(4)} \rightarrow \frac{\tilde{R}}{\sqrt{\tau_2}} \left( Z_1^{(\text{base})} + Z_2^{(\text{base})} \right) \\
\tilde{R} \rightarrow 0: \quad & Z_{\mathcal{X}}^{(4)} \rightarrow \frac{1}{2\sqrt{\tau_2}\tilde{R}} \left( Z_1^{(\text{base})} + Z_3^{(\text{base})} \right). \quad (3.57)
\end{aligned}$$

We thus see that our original four-dimensional theory with partition function  $Z_{\mathcal{X}}^{(4)}$  flows to *different* theories in the  $\tilde{R} \rightarrow \infty$  and  $\tilde{R} \rightarrow 0$  limits. Indeed, from Eq. (3.11) and identifying  $\mathcal{M}^{1/2} V_{1/2}$  as  $\tilde{R}/2$ , we find

$$\begin{aligned}
\tilde{R} \rightarrow \infty: \quad & Z_{\mathcal{X}}^{(5)} = \frac{2}{\sqrt{\tau_2}} (Z_1^{(\text{base})} + Z_2^{(\text{base})}) \\
&= Z_+^+ + Z_+^- + Z_-^+ + Z_-^- \\
\tilde{R} \rightarrow 0: \quad & Z_{\mathcal{X}}^{(5)} = \frac{2}{\sqrt{\tau_2}} (Z_1^{(\text{base})} + Z_3^{(\text{base})}) \\
&= Z_+^+. \quad (3.58)
\end{aligned}$$

We thus see that  $Z_{\mathcal{X}}^{(4)}$  flows to the original five-dimensional *untwisted* theory in the  $\tilde{R} \rightarrow \infty$  limit, while it flows to the five-dimensional *twisted* theory in the  $\tilde{R} \rightarrow 0$  limit. In other words, our four-dimensional theory can be viewed as *interpolating* between two different five-dimensional theories. This kind of interpolation between different decompactified theories is completely standard, and the breaking of  $T$ -duality in this case is the effect of the twist in the compactification.

Our discussion thus far has centered around theories with one large extra dimension compactified on  $S^1/Z_2$ . However, similar treatments will also apply to more complicated compactifications from higher dimensions. For example, it is possible to consider the compactification of a *six*-dimensional theory on a *two*-dimensional compactification geometry. In order to exploit the above results, we can consider this compactification geometry to be  $(S^1/Z_2)_{\tilde{R}_5} \otimes (S^1/Z_2)_{\tilde{R}_6}$  where  $\tilde{R}_{5,6}$  are the dimensionless radii associated with the fifth and sixth dimensions respectively. Our four-dimensional partition function will then have *sixteen* sectors and takes the form

$$Z_{\mathcal{X}}^{(4)} = \sum_{p=1}^4 \sum_{q=1}^4 Z_{pq}^{(\text{base})} Z_{\text{KK/winding}}^{(pq)} \quad (3.59)$$

where  $p$  and  $q$  respectively correspond to the fifth and sixth dimensions and where

$$\begin{aligned}
Z_{\text{KK/winding}}^{(1,1)} &= \tau_2^{-1} \mathcal{E}_0 \cdot \mathcal{E}_0 \\
Z_{\text{KK/winding}}^{(1,2)} &= \tau_2^{-1} \mathcal{E}_0 \cdot \mathcal{E}_{1/2} \\
&\vdots \\
Z_{\text{KK/winding}}^{(4,4)} &= \tau_2^{-1} \mathcal{O}_{1/2} \cdot \mathcal{O}_{1/2}. \quad (3.60)
\end{aligned}$$

Defining

$$g'_{pq}(\tau_2) \equiv \int_{-1/2}^{1/2} d\tau_1 Z_{pq}^{(\text{base})} \quad (3.61)$$

and further defining  $S'_{pq}(\tau_2) = \tau_2 g'_{pq}(\tau_2)$ , we may expand

$$S'_{pq}(\tau_2) \sim \sum_j C_j^{(pq)} \tau_2^j \quad \text{as } \tau_2 \rightarrow 0. \quad (3.62)$$

These  $C_j^{(pq)}$ -coefficients thus correspond to the  $C_j$  coefficients of the simpler untwisted compactification, except that we now have a different set of  $C_j^{(pq)}$ -coefficients for each base theory in Eq. (3.59), i.e., for each value of  $p$  and  $q$ .

For such a four-dimensional theory, there will now be *four* ways of decompactifying in order to produce a six-dimensional theory. These correspond to taking  $\tilde{R}_5 \rightarrow 0, \infty$

and  $\tilde{R}_6 \rightarrow 0, \infty$ , each in principle yielding a *different* six-dimensional theory. The partition functions of these six-dimensional theories will be different combinations of our sixteen underlying  $Z_{pq}^{(\text{base})}$  functions in Eq. (3.59). However, we observe (just as in the five-dimensional case) that no single base function  $Z_i^{(\text{base})}$  by itself corresponds to a decompactified theory. Indeed, this only happens when there is a single sector—i.e., an untwisted compactification.

We shall assume, as stated above, that each decompactification limit leads to a finite one-loop amplitude. Following our previous discussions for the untwisted case, this means that each limit must *independently* satisfy the same smoothness constraint that we imposed in the case of an untwisted compactification. Thus, for the six-dimensional twisted compactification we have been considering here, there are now *four independent* smoothness constraints that must hold. These limits represent the four different ways in which we might obtain a six-dimensional theory.

To formulate these constraints, we can follow our previous analysis in Eq. (3.15) and establish four distinct sums corresponding to these different decompactification limits:

$$S_Q^{(6)} = \tau_2^2 \int_0^\infty d\tau_1 \left[ \lim_Q Z_{\mathcal{X}}^{(4)} \right] \quad (3.63)$$

where  $Z_{\mathcal{X}}^{(4)}$  is given in Eq. (3.59), where  $Q = 1, \dots, 4$  ranges over the different decompactification limits  $(\tilde{R}_5, \tilde{R}_6) \rightarrow (\infty, \infty), (\infty, 0), (0, \infty),$  and  $(0, 0)$  respectively. Each limit will have its own  $\tau_2$ -expansion. To avoid confusion (assuming the reader is not already sufficiently confused), we shall let  $D'$  denote the coefficients of such an expansion:

$$S_Q^{(6)} \sim \sum_j D_j^{(Q)} \tau_2^j \quad \text{as } \tau_2 \rightarrow 0. \quad (3.64)$$

In general, these four sets of  $D'_j$  coefficients (one for each  $Q$ ) will be distinct from each other, with each corresponding to a distinct fully modular-invariant six-dimensional theory.

Given these coefficients, and given our previous discussion, there will be new constraints on each set of coefficients that corresponds to a decompactification limit yielding a finite higher-dimensional amplitude. For example, if the  $\tilde{R}_5 \rightarrow \infty, \tilde{R}_6 \rightarrow \infty$  limit produces a finite string amplitude, then we learn that

$$D_0^{(1)} = 0, \quad D_1^{(1)} = 0. \quad (3.65)$$

Likewise, if the  $\tilde{R}_5 \rightarrow \infty, \tilde{R}_6 \rightarrow 0$  limit also produces a finite string amplitude, then we also have

$$D_0^{(2)} = 0, \quad D_1^{(2)} = 0, \quad (3.66)$$

and so forth. Such results are the twisted analogs of our theorem in Sec. III C, and the proof of these assertions follows directly from the Rankin-Selberg procedure.

Our goal, of course, is to express these  $D^{(Q)}$ -constraints in terms of the  $C_j^{(pq)}$ -coefficients corresponding to our original four-dimensional partition function in Eq. (3.59). These  $C_j^{(pq)}$ -coefficients are defined in Eq. (3.62). However, using Eq. (3.56), we may immediately relate these two sets of coefficients. For example, we find

$$\begin{aligned} D_j^{(1)} &= C_j^{(1,1)} + C_j^{(1,2)} + C_j^{(2,1)} + C_j^{(2,2)} \\ D_j^{(2)} &= C_j^{(1,1)} + C_j^{(1,3)} + C_j^{(3,1)} + C_j^{(3,3)}, \end{aligned} \quad (3.67)$$

and so forth. We thus find that our complete set of constraints becomes

$$\begin{aligned} C_j^{(1,1)} + C_j^{(1,2)} + C_j^{(2,1)} + C_j^{(2,2)} &= 0 \\ C_j^{(1,1)} + C_j^{(1,3)} + C_j^{(2,1)} + C_j^{(2,3)} &= 0 \\ C_j^{(1,1)} + C_j^{(1,2)} + C_j^{(3,1)} + C_j^{(3,2)} &= 0 \\ C_j^{(1,1)} + C_j^{(1,3)} + C_j^{(3,1)} + C_j^{(3,3)} &= 0 \end{aligned} \quad (3.68)$$

for all  $j \leq 1$ .

In the analogous case of an untwisted compactification, we obtained constraints on the  $C'_j$ -coefficients corresponding to the base theory. By contrast, for a twisted compactification, we see our theorem now yields *multiple* constraint equations. However, each of these constrains only a *linear combination* of the coefficients corresponding to *different* base theories. Moreover, as indicated above, each of these constraint equations holds not only for  $j = 0$  but also for  $j = 1$ . As discussed in Sec. III C, the latter reflects the emergence of the extra dimensions and is required for the consistency with our lower-dimensional theory upon decompactification.

Of course, the constraints in Eq. (3.68) allow us to obtain results such as

$$C_j^{(1,2)} + C_j^{(2,2)} = C_j^{(1,3)} + C_j^{(2,3)} \quad (3.69)$$

which do not correspond to any single decompactification limit. Moreover, our four-dimensional theory may also have other internal symmetries that are reflected in constraints on these  $C_j^{(pq)}$ -coefficients. For example, if  $Z_{pq}^{(\text{base})} = Z_{qp}^{(\text{base})}$  (as might occur for theories with a permutation symmetry between the fifth and sixth dimensions), it then follows that  $C_j^{(pq)} = C_j^{(qp)}$  for all  $j$ . In such cases, there are effectively fewer base partition functions and potentially fewer distinct decompactifications as well.

In general, there can also be decompactification limits which are tachyonic. For example, a four-dimensional

theory might be finite (tachyon-free) for compactification volumes within a certain limited range, yet encounter a tachyon as the compactification volume passes beyond this range. Well-known examples of this occur for the thermal analogs of various one-dimensional compactifications, where we identify the compactification radius as an inverse temperature. Indeed, such theories are finite (or even supersymmetric) at zero temperature, and remain finite as the temperature increases within a limited range, but may ultimately become tachyonic once the temperature exceeds a critical value. Indeed, this is the physics that underlies the so-called Hagedorn transition [31–33], which clearly leads to divergences in the one-loop amplitude. As a result, the  $C'$ -constraints that emerge from this decompactification are valid only within the range of radii in which such tachyons do not appear.

There can also be situations in which no tachyon appears at any compactification radius, but in which certain massive states in the string spectrum become massless at specific compactification radii before becoming massive again (see, e.g., Ref. [33]). The sudden appearance of such massless states will generally induce *higher-order* Hagedorn-like phase transitions [33] which represent discontinuities that also violate our “smoothness” assumptions. However, even though our theorem will not apply at or beyond such radii, the constraints emerging from our theorem will continue to apply before these states are reached.

Finally, it is interesting to note that our compactification functions  $\mathcal{E}_{0,1/2}$  and  $\mathcal{O}_{0,1/2}$ —like *any* compactification functions—have certain properties which guarantee that we can continue to use the Rankin-Selberg mapping. In particular, *a priori*, one might have worried that additional  $C$ -constraints could appear upon compactification. It is easy to see how such extra constraints might have arisen. For this purpose, it is perhaps easiest to start with the compactified theory with the partition function given in Eq. (3.55) and ask what happens for large but finite  $R$ . In this regime the terms involving  $\mathcal{E}$ -functions dominate—terms which we can rewrite in the form

$$Z_{\mathcal{X}}^{(4)} \approx \frac{1}{2} \tau_2^{-1/2} \left[ (Z_1^{(\text{base})} + Z_2^{(\text{base})})(\mathcal{E}_0 + \mathcal{E}_{1/2}) + (Z_1^{(\text{base})} - Z_2^{(\text{base})})(\mathcal{E}_0 - \mathcal{E}_{1/2}) \right]. \quad (3.70)$$

The top line, of course, is entirely expected and does not yield any constraints beyond those we have already considered. Indeed,  $Z_1^{(\text{base})} + Z_2^{(\text{base})}$  is the  $R \rightarrow \infty$  limit that we have already considered, and whose properties as  $\tau_2 \rightarrow 0$  have allowed us to formulate our existing constraints. However, in principle, there is the possibility that the *second* line of Eq. (3.70) might lead to additional constraints. Indeed, such additional constraints could have arisen if  $\mathcal{E}_0 - \mathcal{E}_{1/2}$  were for example to behave as  $\tau_2/R$  as  $\tau_2 \rightarrow 0$  at finite  $R$ . In such cases, one could take  $R \rightarrow \infty$

prior to taking  $\tau_2$  and thereby conclude that no new constraint comes from such a difference. However, we could alternatively take the  $\tau_2 \rightarrow 0$  limit first, leaving us with a dangerous  $1/R$  dependence whose cancellation would require an additional constraint.

However, it is straightforward to verify that this does not happen. Indeed, direct calculation for our specific  $\mathcal{E}$ -functions tells us that

$$\mathcal{E}_0(\tilde{R}) - \mathcal{E}_{1/2}(\tilde{R}) \approx (\tilde{R}^4/\tau_2) e^{-\pi\tilde{R}^2/\tau_2}. \quad (3.71)$$

Thus the *difference* between the  $\mathcal{E}$  functions decreases faster than any power of  $\tau_2$  as  $\tau_2 \rightarrow 0$ . In other words, the difference  $\mathcal{E}_0 - \mathcal{E}_{1/2}$  has an essential singularity at  $\tau_2 = 0$ . As a matter of principle this will be true for any compactification functions whose large-volume limits produce valid higher-dimensional theories. Thus our theorem remains intact.

The discussion in this section has been somewhat technical. However, the main idea is simple and can be summarized as follows. For an untwisted compactification, there is generally only one decompactification limit. This then leads to a single extra  $C'$ -constraint, namely  $C'_j = 0$  for  $j < 1 + \delta/2$ , which must be adjoined onto our original  $C$ -constraint  $C_0 = 0$ . However, for a twisted compactification, there are generally multiple decompactification limits, each involving a different subset of the sectors in our theory. Our theorem nevertheless applies exactly as before, with each decompactification leading to its own new  $C'$ -constraint, as illustrated above. As a result, our original four-dimensional theory must satisfy not only the original constraint  $C_0 = 0$  but also all of the individual  $C'$ -constraints that emerge from each different decompactification limit. Indeed, all of these extra  $C'$ -constraints must be satisfied *simultaneously* within the original four-dimensional theory because this theory simultaneously contains within itself all of these different possibilities for a self-consistent decompactification.

We close with a final comment. As discussed in Sec. III D for the case of an untwisted compactification, we have seen that the single extra constraint  $C'_j = 0$  for  $j < 1 + \delta/2$  actually *implied* our original  $C$ -constraint  $C_0 = 0$ . The same will be true for twisted compactifications. In particular, even though each different decompactification leads to an independent  $C'$ -constraint, these  $C'$ -constraints collectively imply that  $C_0 = 0$  as well.

#### IV. IMPLICATIONS OF THE THEOREM

As discussed in Sec. III, our theorem is completely general, providing us with new constraints on the  $C'_j$ -coefficients for any string model meeting the conditions outlined at the beginning of Sec. III A and for any operator insertion  $\mathcal{X}$  for which the corresponding physical quantity  $\zeta \sim \langle \mathcal{X} \rangle^{(4)}$  within that model is finite. In this section we now proceed to consider two important implications of

these new  $C'$ -constraints. For concreteness and simplicity, we shall focus in this section on the case of untwisted compactifications. As discussed in Sec. III, the situation with twisted compactifications is similar and proceeds in an analogous way.

### A. New supertrace identities

Just as the  $C$ -coefficients can be expressed as supertraces over the states within the full partition function  $Z_{\mathcal{X}}^{(4)}$ , as in Eq. (2.34), the  $C'$ -coefficients can likewise be expressed as supertraces of the states that contribute to  $Z_{\mathcal{X}}^{(4)}$  without including the KK and winding states associated with the compactification under study. Indeed, these are the states that reside within the base theory and contribute to  $Z_{\mathcal{X}}^{(\text{base})}$  alone.

Given this observation, our new constraints on the  $C'_j$ -coefficients immediately yield new constraints for the supertraces over the states contributing solely to  $Z_{\mathcal{X}}^{(\text{base})}$ . In particular, the vanishing of the  $C'_j$ -coefficients for all  $j < 1 + \delta/2$ —as required by our theorem—thus gives rise to the new supertrace identities

$$\text{Str}' D_{\tau_2}^j \mathcal{X} = 0 \quad \text{for all } 0 \leq j \leq \delta/2 \quad (4.1)$$

where the prime on the supertrace indicates that only the states contributing to  $Z_{\mathcal{X}}^{(\text{base})}$  are included and where  $D_{\tau_2}$  is the derivative defined in Eq. (2.31). Likewise, as the compactification  $T$ -volume  $\tilde{V}_T$  becomes large, a similar supertrace formulation applies to the amplitude  $\langle \mathcal{X} \rangle^{(4)}$ . We know, of course, that  $\langle \mathcal{X} \rangle^{(4)}$  is given by  $(\pi/3)C_1$  for all compactification volumes. Indeed,  $C_1$  is generally a complicated function of the compactification volume. However, for  $\tilde{V}_T \gg 1$ , we have seen that our  $T$ -volume scaling rule

allows us to pull out the compactification volume as a single overall factor, leaving us with

$$\langle \mathcal{X} \rangle^{(4)} \approx \frac{\pi}{3} \tilde{V}_T C'_{1+\delta/2}. \quad (4.2)$$

Thus, for  $\tilde{V}_T \gg 1$ , we have

$$\langle \mathcal{X} \rangle^{(4)} \approx \frac{\pi}{3} \frac{1}{(1 + \delta/2)!} \tilde{V}_T \text{Str}' D_{\tau_2}^{1+\delta/2} \mathcal{X}, \quad (4.3)$$

thereby once again yielding results depending on a primed supertrace.

The general results in Eqs. (4.1) and (4.3) yield a host of new supertrace constraints on the spectrum of the base theory. In general, the supertrace relations that emerge depend on the number  $\delta$  of extra spacetime dimensions which are associated with our decompactification limits. Indeed, taking  $\mathcal{X}$  of the form in Eq. (2.8), we may extract these new identities directly from Eq. (3.17).

Our results are as follows. Defining

$$\tilde{M}^2 \equiv \frac{M^2}{4\pi\mathcal{M}^2} \quad (4.4)$$

and recalling Eq. (2.35), we see that our usual four-dimensional constraints  $C_0 = 0$  and  $\langle \mathcal{X} \rangle^{(4)} = (\pi/3)C_1$  now take the form

$$\begin{cases} \text{Str}' \mathbb{X}_0 = 0 \\ \langle \mathcal{X} \rangle^{(4)} = \frac{\pi}{3} [\text{Str}' \mathbb{X}_1 - \text{Str}'(\mathbb{X}_0 \tilde{M}^2)]. \end{cases} \quad (4.5)$$

However, our theorem now tells us that for each decompactification limit we have the additional constraints

$$\begin{aligned} \delta = 2: & \begin{cases} \text{Str}' \mathbb{X}_0 = 0 \\ \text{Str}' \mathbb{X}_1 - \text{Str}'(\mathbb{X}_0 \tilde{M}^2) = 0 \\ \langle \mathcal{X} \rangle^{(4)} \approx \frac{\pi}{3} \tilde{V}_T \left[ \text{Str}' \mathbb{X}_2 - \text{Str}'(\mathbb{X}_1 \tilde{M}^2) + \frac{1}{2} \text{Str}'(\mathbb{X}_0 \tilde{M}^4) \right] \end{cases} \\ \delta = 4: & \begin{cases} \text{Str}' \mathbb{X}_0 = 0 \\ \text{Str}' \mathbb{X}_1 - \text{Str}'(\mathbb{X}_0 \tilde{M}^2) = 0 \\ \text{Str}' \mathbb{X}_2 - \text{Str}'(\mathbb{X}_1 \tilde{M}^2) + \frac{1}{2} \text{Str}'(\mathbb{X}_0 \tilde{M}^4) = 0 \\ \langle \mathcal{X} \rangle^{(4)} \approx \frac{\pi}{3} \tilde{V}_T \left[ -\text{Str}'(\mathbb{X}_2 \tilde{M}^2) + \frac{1}{2} \text{Str}'(\mathbb{X}_1 \tilde{M}^4) - \frac{1}{6} \text{Str}'(\mathbb{X}_0 \tilde{M}^6) \right] \end{cases} \\ \delta = 6: & \begin{cases} \text{Str}' \mathbb{X}_0 = 0 \\ \text{Str}' \mathbb{X}_1 - \text{Str}'(\mathbb{X}_0 \tilde{M}^2) = 0 \\ \text{Str}' \mathbb{X}_2 - \text{Str}'(\mathbb{X}_1 \tilde{M}^2) + \frac{1}{2} \text{Str}'(\mathbb{X}_0 \tilde{M}^4) = 0 \\ \text{Str}'(\mathbb{X}_2 \tilde{M}^2) - \frac{1}{2} \text{Str}'(\mathbb{X}_1 \tilde{M}^4) + \frac{1}{6} \text{Str}'(\mathbb{X}_0 \tilde{M}^6) = 0 \\ \langle \mathcal{X} \rangle^{(4)} \approx \frac{\pi}{6} \tilde{V}_T \left[ \text{Str}'(\mathbb{X}_2 \tilde{M}^6) - \frac{1}{3} \text{Str}'(\mathbb{X}_1 \tilde{M}^6) + \frac{1}{12} \text{Str}'(\mathbb{X}_0 \tilde{M}^8) \right], \end{cases} \end{aligned} \quad (4.6)$$

where our expressions for  $\langle \mathcal{X} \rangle^{(4)}$  in each case are appropriate for  $\tilde{V}_T \gg 1$ . Indeed, given these results, we may regard the results in Eq. (4.5) as “ $\delta = 0$ ” constraints, with the understanding that  $\text{Str}' \rightarrow \text{Str}$  in this case. We shall continue to refer to our original four-dimensional results as  $\delta = 0$  results in the following.

As we shall discuss below, the supertrace identities in Eq. (4.6) for each value of  $\delta$  hold exactly regardless of whether Eq. (3.1) is satisfied. By contrast, our expressions for the amplitude  $\langle \mathcal{X} \rangle^{(4)}$  hold only when Eq. (3.1) is satisfied. However, even when Eq. (3.1) is not satisfied, we will find that these expressions for  $\langle \mathcal{X} \rangle^{(4)}$  continue to hold except that an additional term is added in each case.

Note that the exact value for  $\langle \mathcal{X} \rangle^{(4)}$  for *any* compactification volume is given in Eq. (4.5). The correctness of this result does not change even as the compactification volume becomes large. However, the behavior of this amplitude for large compactification volume is not readily apparent from the expression in Eq. (4.5). By contrast, in Eq. (4.6), we have provided expressions for  $\langle \mathcal{X} \rangle^{(4)}$  which approximate this true value for  $\tilde{V}_T \gg 1$  and in which the compactification  $T$ -volume  $\tilde{V}_T$  emerges simply as an overall multiplicative factor.

These results are completely general, written in terms of the arbitrary  $\mathbb{X}_\ell$  insertions. However, using Eqs. (2.10) and (2.18), we can write these expressions directly in terms of the operators relevant for the Higgs mass, as in Ref. [13], or the one-loop gauge coupling, as in Ref. [17].

In the case of the Higgs mass, there is actually an important short-cut that we may exploit. As shown in Ref. [13], one can write the one-loop Higgs mass as

$$m_\phi^2 = \left( \partial_\phi^2 + \frac{\xi}{4\pi^2} \right) \Lambda(\phi) \Big|_{\phi=0} \quad (4.7)$$

where  $\phi$  denotes a fluctuation of the Higgs field relative to its VEV within the Higgsed phase of any string model, where  $\xi$  is a model-dependent numerical parameter defined in Eq. (2.15), and where  $\Lambda(\phi)$  is the amplitude corresponding to the trivial insertions

$$\mathbb{X}_0 = -\frac{1}{2} \mathcal{M}^4, \quad \mathbb{X}_1 = \mathbb{X}_2 = 0 \quad (4.8)$$

where the masses of the states in the string spectrum are generally treated as functions of  $\phi$ . It is the choice of such functions which specifies the particular scalar field whose mass is being calculated. Indeed, thinking of  $\Lambda$  as a kind of  $\phi$ -dependent Coleman-Weinberg potential, we thus see from Eq. (4.7) that the Higgs mass is essentially given by the curvature of  $\Lambda(\phi)$  around its minimum, augmented [13] by a stringy gravitational backreaction

parametrized by  $\xi$ . Indeed,  $\Lambda \equiv \Lambda(\phi)|_{\phi=0}$  is nothing but the one-loop cosmological constant.

Thus, first performing our analysis for  $\Lambda(\phi)$ , we find that the constraints in Eq. (4.1) now take the simple form [2]

$$\text{Str} M^{2k}(\phi) = 0 \quad \text{for all } k \leq \delta/2. \quad (4.9)$$

Likewise Eq. (4.2) yields

$$\Lambda(\phi) \approx \mathcal{M}^{-\delta} \tilde{V}_T \Lambda^{(4+\delta)}(\phi) \quad (4.10)$$

where the higher-dimensional cosmological constant  $\Lambda^{(4+\delta)}$  follows from Eq. (4.3) and is given by [2]

$$\Lambda^{(4+\delta)}(\phi) = \frac{\pi}{3} \frac{\mathcal{M}^{4+\delta}}{2} \frac{(-1)^{\delta/2}}{(1+\delta/2)!} \text{Str}'[(\tilde{M}^2)^{1+\delta/2}] \quad (4.11)$$

where we also regard  $\tilde{M}^2$  as a  $\phi$ -dependent quantity. We observe, in this context, that the  $k = 0$  equation within Eq. (4.9) is nothing but the constraint [2]

$$\text{Str} \mathbf{1} = 0. \quad (4.12)$$

This result, which actually applies to *all* finite four-dimensional theories regardless of whether they have decompactification limits, implies that all such theories have equal numbers of bosonic and fermionic degrees of freedom *when summed across the entire string spectrum*. Indeed, this observation holds even through there is no boson/fermion pairing, and is a nontrivial result of the UV/IR mixing inherent in such theories in which any surpluses of bosonic or fermionic degrees of freedom of a given mass are balanced against opposite surpluses at other masses throughout the string spectrum. Likewise, the exact result for the cosmological constant in four dimensions, specifically [2]

$$\Lambda = \frac{1}{24} \mathcal{M}^2 \text{Str} M^2, \quad (4.13)$$

is equally surprising, telling us that the full one-loop zero-point amplitude in such theories is given simply as a supertrace of the squared masses of the string states across the string spectrum, with the regulator within the definition of the supertrace in Eq. (2.33) ensuring a finite result.

Given these results for  $\Lambda(\phi)$ , we can now apply the result in Eq. (4.7) in order to obtain our corresponding results for the Higgs mass  $m_\phi^2$ . In particular, for  $\delta = 0$  and  $\delta = 2$  we obtain

$$\delta = 0: \begin{cases} \text{Str}\mathbf{1} = 0 \\ m_\phi^2 = \frac{\mathcal{M}^2}{24} \left[ \text{Str}(\partial_\phi^2 M^2) + \frac{\xi}{4\pi^2 \mathcal{M}^2} \text{Str} M^2 \right] \Big|_{\phi=0} \end{cases}$$

$$\delta = 2: \begin{cases} \text{Str}'\mathbf{1} = 0 \\ \text{Str}'\partial_\phi^2 M^2 = 0 \\ m_\phi^2 \approx -\frac{1}{48\pi} \tilde{V}_T \left[ \text{Str}'(\partial_\phi^2 M^4) + \frac{\xi}{4\pi^2 \mathcal{M}^2} \text{Str}' M^4 \right] \Big|_{\phi=0}, \end{cases} \quad (4.14)$$

and so forth for higher  $\delta$ . In this connection, we note that the second constraint equation for  $\delta = 2$  actually takes the full form

$$\text{Str}'(\partial_\phi^2 M^2) + \frac{\xi}{4\pi^2} \text{Str}' M^2 = 0. \quad (4.15)$$

However, we see from Eq. (4.9) that the second term actually vanishes, thus leaving us with the simpler constraint equation in Eq. (4.14). Similar cancellations would likewise occur for  $\delta > 2$  involving  $\text{Str} M^{2k}$  for  $k > 1$ . Likewise, we emphasize that the *constraint* equations in Eq. (4.14) actually hold for all  $\phi$ , whereas our expressions for  $m_\phi^2$  hold by definition only for  $\phi = 0$ . Finally, we note that the results in Eq. (4.14)—or more specifically, their truncations to  $\phi = 0$ —could have been obtained directly using the operator insertions in Eq. (2.10). This provides an important cross-check on our calculations.

Before proceeding, we note that the constraint equations in Eq. (4.9) are highly nontrivial. When truncated to  $\phi = 0$ , these results provide tight constraints on the masses of all of the states in our theory. Indeed, for  $\delta = 6$ , we learn that the states throughout the string spectrum must, through UV/IR

mixing, arrange themselves so as to simultaneously cancel not only  $\text{Str}\mathbf{1}$ , as in Eq. (4.12), but also  $\text{Str}' M^2$ ,  $\text{Str}' M^4$ , and  $\text{Str}' M^6$ . However, Eq. (4.9) actually holds *as a function of  $\phi$*  for all  $\phi$ . Therefore these equations also constrain how  $\phi$  couples to all of the states in the string spectrum (and not just the massless states):  $\phi$  can only couple in a way that maintains these cancellations as  $\phi$  varies. This thereby restricts which kinds of Higgs fields are ultimately allowed in the theory. Likewise, for any Higgs field, we may also view this as providing a significant constraint on the kinds of fluctuations which are ultimately permitted by the modular invariance of the underlying theory.

Let us now turn to the case of the one-loop contributions to the gauge couplings. As in Ref. [17], these one-loop contributions to the quantity  $16\pi^2/g_G^2$  will be denoted  $\Delta_G$  where  $G$  is the corresponding gauge group. Note in this context that  $\Delta_G$  is the full one-loop contribution to  $16\pi^2/g_G^2$ , and not merely the contribution from the infinite towers of massive states.

Once again, just as with the Higgs mass, we may consider the cases with  $\delta = 0$ ,  $\delta = 2$ , and  $\delta = 4$ . Recalling that  $\mathbb{X}_0 = 0$  for this calculation, we find

$$\delta = 0: \quad \Delta_G = \frac{\xi}{6} \left[ \text{Str} \bar{Q}_H^2 - \frac{1}{12} \text{Str}_E \mathbf{1} \right]$$

$$\delta = 2: \quad \begin{cases} \text{Str}' \bar{Q}_H^2 - \frac{1}{12} \text{Str}_E' \mathbf{1} = 0 \\ \Delta_G \approx \frac{\pi}{3} \tilde{V}_T \left[ -2 \text{Str}'(Q_G^2 \bar{Q}_H^2) + \frac{1}{6} \text{Str}_E' Q_G^2 - \frac{\xi}{2\pi} \text{Str}'(\bar{Q}_H^2 \tilde{M}^2) + \frac{\xi}{24\pi} \text{Str}_E' \tilde{M}^2 \right] \end{cases}$$

$$\delta = 4: \quad \begin{cases} \text{Str}' \bar{Q}_H^2 - \frac{1}{12} \text{Str}_E' \mathbf{1} = 0 \\ -2 \text{Str}'(Q_G^2 \bar{Q}_H^2) + \frac{1}{6} \text{Str}_E' Q_G^2 - \frac{\xi}{2\pi} \text{Str}'(\bar{Q}_H^2 \tilde{M}^2) + \frac{\xi}{24\pi} \text{Str}_E' \tilde{M}^2 = 0 \\ \Delta_G \approx \frac{\pi}{3} \tilde{V}_T \left[ 2 \text{Str}'(\bar{Q}_H^2 Q_G^2 \tilde{M}^2) - \frac{1}{6} \text{Str}_E'(Q_G^2 \tilde{M}^2) + \frac{\xi}{4\pi} \text{Str}'(\bar{Q}_H^2 \tilde{M}^4) - \frac{\xi}{48\pi} \text{Str}_E' \tilde{M}^4 \right]. \end{cases} \quad (4.16)$$

We again emphasize that these supertrace identities involve the full infinite towers of states in our theories. In general, these identities are not satisfied level-by-level, but rather represent conspiracies across the infinite string spectrum at all energy scales. Indeed, these identities are satisfied through the hidden so-called “misaligned supersymmetry” that remains in the string spectrum even after

spacetime supersymmetry is broken in any modular-invariant tachyon-free string theory [1,2].

We note that we have now collected several combinations of supertraces that have to vanish exactly. One might imagine that these particular linear sums of supertraces could cancel between themselves in unique modular-invariant combinations. However *any* modular-invariant

integral that we can construct which we know to be finite provides its own vanishing supertrace constraint. Thus we are at liberty to consider constraints from additional integrals that we also know to be finite even if they do not correspond to any physical process. In Ref. [34] we play this game and derive even stronger sets of supertrace constraints.

### B. Implications for one-loop running: IR limit

Our theorem also has ramifications for the effective field theories (EFTs) that might be associated with our string theories, and in particular the manner in which these EFTs evolve as we change the relevant energy scale  $\mu$  at which they are probed.

To study this, let us first briefly recall how one can extract an EFT from a given string theory in a manner that is consistent with modular invariance and which naturally takes the full spectrum of the string theory into account. Certain aspects of this procedure were discussed previously in Refs. [13,17], where more details can be found.

First, for any four-dimensional amplitude  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$ , we introduce a regulator. Then, just as for ordinary EFTs, we identify the regulator parameter with a running physical scale  $\mu$ . To carry out this procedure we will adopt the same regulator function  $\mathcal{G} \equiv \mathcal{G}_\rho(a, \tau)$  (previously denoted  $\hat{\mathcal{G}}_\rho(a, \tau)$ ) utilized in Refs. [13,17], namely

$$\mathcal{G}_\rho(a, \tau) = \frac{a^2}{1 + \rho a^2} \frac{\rho}{\rho - 1} \frac{\partial}{\partial a} [Z_{\text{circ}}(\rho a) - Z_{\text{circ}}(a)] \quad (4.17)$$

where  $Z_{\text{circ}}$  is the circle-compactification partition function in Eq. (3.5) and where  $\rho$  is a constant which in effect plays the role of an RG scheme. It then follows that our regulated amplitude, which we shall denote  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$ , is given by

$$\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)} = \frac{a^2}{1 + \rho a^2} \frac{\rho}{\rho - 1} \frac{\partial}{\partial a} [P(\rho a) - P(a)] \quad (4.18)$$

where the “reduced” amplitude  $P(a)$  is given by

$$P(a) = \langle \mathcal{X} Z_{\text{circ}}(a, \tau) \rangle. \quad (4.19)$$

Having adopted this convention, we shall then make the identification [13,17]

$$\alpha' \mu^2 \equiv \rho a^2. \quad (4.20)$$

It then follows that our regulated physical quantities such as  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$  are functions of  $\mu$ .

In general, this choice of regulator function  $\mathcal{G}_\rho(a, \tau)$  vanishes exponentially quickly as  $\tau_2 \rightarrow \infty$  (thereby ensuring the effectiveness of this function as an IR regulator) but otherwise asymptotes to 1 as  $\tau_2 \rightarrow 1$  (thereby preserving the original theory in this regime). Indeed, a rough measure

of the transition between these two behaviors occurs at  $\tau_2 \approx (\rho a^2)^{-1}$ . Thus  $\rho a^2$  sets the “scale” at which a given state with mass  $M$  is either included amongst or excluded from (or “integrated out” from) the dynamical degrees of freedom in our analysis. This allows us to make the identification in Eq. (4.20). We also point out that while this identification procedure for  $\mu$  is, strictly speaking, valid only for  $\mu \ll M_s$ , we shall treat this as a *definition* of  $\mu$  within the entire range  $0 \leq \mu \leq M_s$ .

In the following, without loss of generality, we shall adopt the choice  $\rho = 2$ . Given these conventions, our goal is to understand how our theorem affects the *running* of  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$  for arbitrary insertion  $\mathcal{X}$ —i.e., affects the manner in which  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$  varies with  $\mu$ . Our goal is also to understand the corresponding beta-function

$$\beta_{\mathcal{X}}(\mu) \equiv \frac{\partial \langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}}{\partial \log \mu}. \quad (4.21)$$

In general, the result for  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$  for arbitrary  $\mathcal{X}$  and arbitrary  $\mu$  was derived in Refs. [13,17]. This result will also be quoted in Sec. V, where explicit examples will be considered. However, schematically, in cases for which  $\text{Str} \mathbb{X}_2 = 0$  this result takes the form [17]

$$\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu) = \langle \mathcal{X} \rangle^{(4)} + \sum_i c_i \mathcal{K}(M_i/\mu) \quad (4.22)$$

where  $\sum_i$  denotes a sum over the on-shell states in the theory and where  $\mathcal{K}(x)$  schematically denotes a generic combination of Bessel functions. Although a slight generalization of Eq. (4.22) applies when  $\text{Str} \mathbb{X}_2 \neq 0$ , as discussed in Refs. [13,17], we shall continue to assume that  $\text{Str} \mathbb{X}_2 = 0$  throughout this discussion.

Let us begin by studying the behavior of  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$  for extreme values of  $\mu$ , i.e., for  $\mu = 0$  as well as  $\mu = M_s$ . Having fixed the behavior of  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}$  at these extreme endpoints, we shall then investigate the implications of our theorem for the intermediate values  $0 < \mu < M_s$ .

We first discuss the IR behavior of  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$ , i.e., the behavior of this amplitude as  $\mu \rightarrow 0$ . This behavior can be extracted [17] from the asymptotic properties of the Bessel function combinations  $\mathcal{K}$ , yielding

$$\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu) \approx \langle \mathcal{X} \rangle^{(4)} + \sum_{0 < M \leq \mu} b_M \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E + 1) \frac{\mu}{M}} \right], \quad (4.23)$$

where

$$b_M \equiv -2 \text{Str}_{\text{mass}=M} \mathbb{X}_2, \quad (4.24)$$

and where  $\gamma_E \approx 0.57721$  is the Euler-Mascheroni constant. Note that the supertrace in Eq. (4.24) is over only those string states whose masses are equal to  $M$ .

Having introduced the regulator, we are now able to relax the condition in Eq. (3.1) and also consider the additional contributions from massless  $\mathbb{X}_2$ -charged states, i.e., the additional contributions that would arise if  $\text{Str}_{M=0} \mathbb{X}_2 \neq 0$ .

Following the algebraic steps in Ref. [13], it turns out that our result in Eq. (4.23) would accrue the additional term

$$b_G \log \left[ \frac{e^{-1} \mu}{2\sqrt{2} M_s} \right] \quad (4.25)$$

where

$$b_G \equiv -2 \text{Str}_{M=0} \mathbb{X}_2. \quad (4.26)$$

Note that while the expression in Eq. (4.23) depends on the logarithm of  $\mu/M$ , the expression in Eq. (4.25) depends on the logarithm of  $\mu/M_s$ . This difference has an important effect. Within Eq. (4.23) we are instructed to consider only those values of  $\mu$  which exceed  $M$ . Thus this contribution remains finite for all  $\mu$ , given that  $M > 0$ . However, within Eq. (4.25), we are permitted to consider values of  $\mu$  into the deep IR. In such cases the logarithm within Eq. (4.25) diverges. However, this divergence is expected, as this is nothing but the divergence that is caused by having  $\text{Str}_{M=0} \mathbb{X}_2 \neq 0$ .

### C. Generalized $T$ -volume scaling rule

The use of our regulator also allows us to generalize the  $T$ -volume scaling rule in Eq. (4.2) which states that  $\langle \mathcal{X} \rangle^{(4)} \approx (\pi/3) \tilde{V}_T C'_{1+\delta/2}$  for large compactification volume. Of course, this result also assumed that  $b_G = 0$ . However, when  $b_G \neq 0$ , we may easily derive a more general result. To do this, we begin with our regulated amplitude in Eq. (2.20), specifically

$$\begin{aligned} \langle \mathcal{X} \rangle^{(4)} &= \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} Z_{\mathcal{X}}^{(\text{base})} Z_{\text{KK/winding}} \mathcal{G}_\rho(a, \tau) \\ &= \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} [\tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})}] [\tau_2^{\delta/2} Z_{\text{KK/winding}}] \\ &\quad \times \mathcal{G}_\rho(a, \tau) \end{aligned} \quad (4.27)$$

where as in Eq. (3.7) we have reshuffled the factors into modular-invariant blocks, and where for generality we have kept  $\delta$  arbitrary. In order to separate out the massless  $\mathbb{X}_2$ -charged contribution from such an expression we need only realize that  $\text{Str}_{M=0} \mathbb{X}_2$  is itself a modular-invariant quantity. This follows from the fact that in the absence of a regular this quantity serves as the coefficient of

the logarithmic divergence. We may therefore rewrite Eq. (4.27) in the form

$$\begin{aligned} \langle \mathcal{X} \rangle^{(4)} &= \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} [\tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})} - b_G] [\tau_2^{\delta/2} Z_{\text{KK/winding}}] \mathcal{G}_\rho(a, \tau) \\ &\quad + b_G \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} [\tau_2^{\delta/2} Z_{\text{KK/winding}}] \mathcal{G}_\rho(a, \tau). \end{aligned} \quad (4.28)$$

Both of these integrals are now explicitly finite and also modular invariant. Moreover, the first of these integrals would have been finite (because we have subtracted  $b_G$  within the first integrand) even if we were to neglect the regulator entirely. Stated differently, the quantity  $\tau_2^{-\delta/2} Z_{\mathcal{X}}^{(\text{base})} - b_G$  rapidly decays as  $\tau_2 \rightarrow \infty$ , and thus receives the bulk of its support from values of  $\tau_2$  at which the regulator  $\mathcal{G}_\rho(a, \tau)$  is close to unity, by construction. We can therefore disregard the appearance of  $\mathcal{G}_\rho(a, \tau)$  within the first integrand, whereupon the previous analysis in Sec. III E allows us to write the first integral as  $(\pi/3) \tilde{V}_T (C'_{1+\delta/2} - b_G)$ . We thus obtain the result

$$\begin{aligned} \langle \mathcal{X} \rangle^{(4)} &\approx \frac{\pi}{3} \tilde{V}_T (C'_{1+\delta/2} - b_G) \\ &\quad + b_G \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} \tau_2^{\delta/2} Z_{\text{KK/winding}} \mathcal{G}_\rho(a, \tau) \end{aligned} \quad (4.29)$$

which generalizes our original  $T$ -volume scaling rule to cases in which  $b_G \neq 0$ . Here  $C'_{1+\delta/2}$  is the supertrace over the entire spectrum of the base theory, and  $\tilde{V}_T$  continues to be given in Eq. (3.39). This result thus allows us to extract the IR behavior of our string-theoretic amplitude.

### D. Implications for one-loop running: UV limit and scale duality

Next, we discuss the behavior which arises in the extreme UV limit as  $\mu \approx M_s$ . As discussed in Refs. [13, 17], our result for  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$  generally exhibits a *scale-duality* symmetry under which  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$  is invariant under

$$\mu \rightarrow \frac{M_s^2}{\mu}. \quad (4.30)$$

Indeed, this symmetry applies to all such amplitudes regardless of the particular operator insertion  $\mathcal{X}$ . However, due to this symmetry and the fact that  $\langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)$  is a smooth function of  $\mu$ , we learn that

$$\left. \frac{d \langle \mathcal{X} \rangle_{\mathcal{G}}^{(4)}(\mu)}{d\mu} \right|_{\mu=M_s} = 0. \quad (4.31)$$

In other words, the corresponding  $\beta$ -function  $\beta_{\mathcal{X}}(\mu)$  actually vanishes at the string scale, implying that there is an apparent *UV fixed-point* regime around  $\mu = M_s$ . As might be expected, this is a purely stringy effect which cannot be captured through an EFT-based analysis.

### E. Absence of power-law running

Having discussed the behavior of  $\langle \mathcal{X} \rangle_G^{(4)}(\mu)$  for  $\mu \approx 0$  and  $\mu \approx M_s$ , we now investigate the behavior of the *running* between these two endpoints. It is here that our theorem leads to some additional surprising effects, connecting the two extremes  $\mu \rightarrow 0$  and  $\mu \approx M_s$  in sometimes unexpected ways.

In general, for large-volume compactifications, it is a natural (ultimately field-theoretic) expectation that at energy scales exceeding the compactification scale, the accumulation of contributions from increasing numbers of Kaluza-Klein states running in loops will slowly deform an expected four-dimensional logarithmic running for a given amplitude into a *power-law* running, as consistent with the emergence of extra spacetime dimensions (or equivalently an increase in the effective dimensionality of the theory) in this limit [35,36]. More specifically, in the case of the one-loop inverse gauge couplings  $\Delta_G(\mu)$ , and at scales  $\mu \gtrsim R^{-1}$  where  $R$  is a large compactification radius, one generally expects the logarithmic running we observe in Eq. (4.23) to follow a power law instead, with  $\Delta_G(\mu) \sim (\mu R)^\delta$ . Indeed, in such situations the volume of the compactification manifold can be taken as  $V_\delta \sim (2\pi R)^\delta$ . Note that we are here referring to  $V_\delta$  as our compactification volume rather than  $\tilde{V}_T$  because we are discussing our field-theoretic expectations.

It is easy to see how such a result might arise from the logarithmic term within our result in Eq. (4.23), specifically the term

$$\Delta_G(\mu) \supset 4 \operatorname{Str}_{0 < M \lesssim \mu} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E+1)} \frac{\mu}{M} \right]. \quad (4.32)$$

To keep the discussion simple, we shall focus on the most straightforward case in which our compactification is *untwisted*, so that each state in the theory with a given value of  $\bar{Q}_H^2$  and  $Q_G^2$  has an infinite spectrum of KK copies with higher masses but the same values of  $\bar{Q}_H^2$  and  $Q_G^2$ . Of course, at any energy scale  $\mu$ , the supertrace is over all states in the theory with masses  $0 < M \lesssim \mu$ . Thus, given this KK structure, our supertrace *factorizes*, i.e.,

$$\operatorname{Str}_{0 < M \lesssim \mu} = \operatorname{Str}'_{0 < M \lesssim \mu} \cdot \operatorname{Tr}_{0 < M \lesssim \mu}^{(\text{KK})}, \quad (4.33)$$

where  $\operatorname{Str}'_{0 < M \lesssim \mu}$  is a supertrace over the different states in the theory *excluding* the excited KK modes associated with the

large dimensions—indeed, the states in this sum may be regarded as the corresponding KK zero modes—and  $\operatorname{Tr}_{0 < M \lesssim \mu}^{(\text{KK})}$  is a trace over the excited KK modes associated with the large compactified dimensions. Since these excited KK states necessarily have the same spins as their corresponding zero modes, the latter is a trace rather than a supertrace. In other words, the  $(-1)^F$  factor has already been absorbed into the primed supertrace rather than the KK trace. We further note that we do not consider the effects of any winding modes, since this is meant to be a purely field-theoretic analysis.

Given the factorization in Eq. (4.33), we can rewrite Eq. (4.32) as

$$\Delta_G(\mu) \supset 4 \operatorname{Str}'_{0 < M \lesssim \mu} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 \times \operatorname{Tr}_{0 < M \lesssim \mu}^{(\text{KK})} \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E+1)} \frac{\mu}{M} \right]. \quad (4.34)$$

Let us now focus for simplicity on the  $\delta = 1$  case in which our theory is compactified on an untwisted circle of radius  $R$ . In such a case, our excited KK spectrum consists of states with masses  $M_{\tilde{m}} \equiv \tilde{m}/R$ , with  $\tilde{m} = \pm 1, \pm 2, \pm 3, \dots$ . For  $\mu R \gg 1$ , we then find that the final line of Eq. (4.34) becomes [35,36]

$$\begin{aligned} & 2 \sum_{\tilde{m}=1}^{\mu R} \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E+1)} \frac{\mu R}{\tilde{m}} \right] \\ &= 2\mu R \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E+1)} \mu R \right] - 2 \log(\mu R)! \\ &\approx 2 \left\{ 1 + \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma_E+1)} \right] \right\} \mu R \end{aligned} \quad (4.35)$$

where we used Stirling's approximation  $\log N! \approx N \log N - N$  for  $N \gg 1$  in passing to the final line. We thus see that the sum over the KK modes associated with a single compactified extra dimension has changed our logarithmic running to a running which is *linear* in  $\mu$ . This process easily generalizes to the KK excitations associated with  $\delta$ -dimensional compactification manifolds with  $\delta > 1$ , yielding power-law running with correspondingly higher powers  $\Delta_G(\mu) \sim (\mu R)^\delta$ .

This phenomenon whereby a sum over KK states deforms a running from logarithmic to power-law is well known from phenomenological studies of theories with large extra dimensions [35,36]. Indeed, it played a crucial role in realizing low-scale gauge-coupling unification [35,36]. As mentioned above, this result can ultimately be understood from the observation that a large compactification radius  $R$  effectively increases the overall spacetime dimensionality of the theory for all energy scales  $\mu \gtrsim R^{-1}$ , thereby shifting the

mass dimensions of the gauge couplings and consequently shifting their corresponding runnings.

Before going further, several remarks are in order. First, the above discussion has assumed an *untwisted* compactification—this is what enabled the complete factorization of the supertraces in Eq. (4.33). However, even for twisted compactifications in which such a complete factorization does not apply, the theory can be separated into individual sectors, and such factorizations are valid within each sector. One then obtains the same power-law results sector by sector. Second, the discussion above has neglected the contributions of winding modes. However, this would appear to be justified because we are performing a field-theoretic analysis, and also because we are restricting our attention to cases with  $\mu \ll M_s$  and  $R^{-1} \ll \mu$ . In general, for large string compactifications with  $R^{-1} \ll \mu$ , the corresponding winding modes will have masses  $\gg \mu$ . Thus—from a field-theoretic perspective—such states will not directly affect the running at the scale  $\mu$ . Finally, we have taken all of our KK masses as  $M = \tilde{m}/R$ . Of course, this is correct for the KK excitations of only the massless states within the primed supertrace. However, for  $\mu < M_s$ , this is a valid assumption because we are assuming that the KK modes associated with any large dimensions are already part of the KK trace, because the winding states have masses above  $M_s$ , and because the only other states—the string excitations—have masses which are at (or heavier than) the string scale. Thus the only states which contribute to the primed supertrace are indeed those which are massless.

There is, however, one important shortcoming to the above treatment: we did not handle the spacetime compactification in a fully modular-invariant manner. Indeed, we simply split our supertrace into separate pieces without considering the deeper aspects of the effects of the compactification on the underlying *partition function*. However, it is at the level of the partition function that modular invariance must be maintained. Although the above treatment captures the expected higher-dimensional power-law running—and would thus be sufficient for a field-theoretic analysis, as in Refs. [35,36]—it misses the critical fact that a fully modular-invariant theory in higher dimensions  $D > 4$  has more internal cancellations within its spectrum than does a four-dimensional theory. As we have seen, these extra internal cancellations are required by modular invariance, and in particular can be attributed to various supertrace identities, such as  $\text{Str} \mathbf{1} = 0$ , that result from UV/IR mixing and misaligned supersymmetry. Indeed the higher the spacetime dimensionality of a modular-invariant theory, the more internal cancellations of this sort exist within the spectrum.

As we shall now demonstrate, our theorem implies that these cancellations ultimately have the effect of *eliminating* the above power-law running. This is a remarkable result, implying that there is actually *no running at all* above the compactification scale.

To understand this, let us begin by recalling that our running four-dimensional amplitude  $\langle \mathcal{X} \rangle^{(4)}$  is given by

$$\langle \mathcal{X} \rangle^{(4)}(\mu) = \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z_{\mathcal{X}}^{(\text{base})} Z_{\text{KK/winding}} \mathcal{G}(\mu, \tau). \quad (4.36)$$

Here  $\mathcal{G}(\mu, \tau)$  is the “regulator” function whose purpose is not to regulate this amplitude (since the amplitude is already presumed finite), but rather to introduce the running scale  $\mu$ . For example, as we have seen, we may identify  $\mathcal{G}(\mu, \tau) = \mathcal{G}_\rho(a, \tau)$  where  $\mathcal{G}_\rho(a, \tau)$  is defined in Eq. (4.17) and where the scale  $\mu$  is defined in terms of  $\rho$  and  $a$  in Eq. (4.20). Ultimately, our claim is that  $\langle \mathcal{X} \rangle^{(4)}$  does not run as a power-law function of  $\mu$ . However, it turns out that this absence of power-law running is wholly independent of the specific form of  $\mathcal{G}(\mu, \tau)$ , and instead relies on deeper, more universal features associated with the modular invariance of the partition function  $Z_{\mathcal{X}}^{(\text{base})}$ .

This can most easily be understood by recasting the above string-theoretic expression into its field-theory analog. This will ultimately enable us to make a direct comparison between the two, and thereby uncover the reason for the absence of power-law running in field-theory language. The field-theory analog of the expression in Eq. (4.36) is given by the Schwinger representation of the analogous field-theory amplitude, and takes the form

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = \int_{\Lambda^{-2}}^{\infty} \frac{dt}{t} Z_{\mathcal{X}}^{(\text{base})} Z_{\text{KK}} \mathcal{G}(\mu, t). \quad (4.37)$$

Here  $t$  is the Schwinger proper time (with a field-theoretic measure  $dt/t$ ), while  $\Lambda$  is an arbitrary cutoff on the Schwinger integral and  $Z_{\text{KK}}$  denotes a trace over *only* the KK modes associated with our  $\delta$  extra dimensions (since the winding modes are intrinsically stringy).

Likewise,  $Z_{\mathcal{X}}^{(\text{base})}$  as before tallies the contributions of the physical states in our theory with the above KK states excluded. Of course, in field theory we also now have a UV divergence. We have therefore introduced a UV cutoff  $\Lambda$  in Eq. (4.37). Finally, in this field-theory expression we are also introducing the running scale  $\mu$  just as we would in string theory, through the introduction of a  $\mathcal{G}$ -function  $\mathcal{G}(\mu, \tau)$  which suppresses the contributions to the integral from the region of integration with  $t \gtrsim \mu^{-2}$ . Once again, as for the string-theory amplitude, the absence of running for this field-theory amplitude will be insensitive to the details of this regulator function. We can therefore choose to model this function in the most simple way possible, namely as providing a hard step-function cutoff:

$$\mathcal{G}(\mu, t) = \Theta(\mu^{-2} - t). \quad (4.38)$$

We then have

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = \int_{\Lambda^{-2}}^{\mu^{-2}} \frac{dt}{t} Z_{\mathcal{X}}^{(\text{base})} Z_{\text{KK}}. \quad (4.39)$$

The next step, as in string theory, will be to expand both  $Z_{\mathcal{X}}^{(\text{base})}$  and the product  $Z_{\mathcal{X}}^{(\text{base})} \cdot Z_{\text{KK}}$  in powers of  $t$  as  $t \rightarrow 0$ . In analogy with our string-theory results, we have

$$\begin{aligned} Z_{\mathcal{X}}^{(\text{base})} &= \frac{1}{t^2} (C'_0 + C'_1 t + C'_2 t^2 + \dots) \\ Z_{\mathcal{X}}^{(\text{base})} \cdot Z_{\text{KK}} &= \frac{1}{t^2} (C_0 + C_1 t + C_2 t^2 + \dots) \end{aligned} \quad (4.40)$$

as  $t \rightarrow 0$ . Note that in the string-theory calculation we would be instead  $\tau_2$ -expanding the  $g(\tau_2)$  functions associated with our partition functions, not the partition functions directly. However, in this analogous field-theory system our Schwinger time  $t$  is the analog of  $\tau_2$ ; in particular, we lack an analog of  $\tau_1$ . Thus, whereas in string theory we had both a partition function  $Z$  and a corresponding  $g$ -function, the former involving both physical and unphysical states and the latter involving only physical states, in our field-theory calculation our partition function is already restricted to physical states. Thus both  $Z$  and  $g(\tau_2)$  have the same field-theory analog, which we may simply regard as  $Z$  itself.

It is important to properly interpret the  $C$ - and  $C'$ -coefficients in Eq. (4.40). If we were to take  $\mathcal{X} = \mathbf{1}$  as an example,  $C_0$  becomes the coefficient of the leading divergence, corresponding to the power of  $1/t^3$  that would appear in a calculation of the standard Coleman-Weinberg (CW) potential. Indeed, this would correspond to the quartic  $\Lambda^4$  divergence of the CW potential. Likewise,  $C_1$  would correspond to the quadratic divergence of the CW potential, and so forth.

For this discussion, however, we are interested in situations in which we have  $\delta$  large extra dimensions of radius  $R$ . Our field-theory expectation is that  $\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu)$  in Eq. (4.39) will experience  $\log \mu$  running for  $\mu \lesssim R^{-1}$ , but that for  $\mu \gg R^{-1}$  this logarithmic running is promoted to power-law running  $\sim (\mu R)^\delta$  due to the presence of  $Z_{\text{KK}}$ .

It is easy to see how these expectations arise. For  $\mu \lesssim R^{-1}$ , we know that this integral must have at most a logarithmic dependence on  $\mu$ . We also know that for  $\mu \lesssim R^{-1}$  only the zero modes of  $Z_{\text{KK}}$  contribute, so that we can approximate  $Z_{\text{KK}} = 1$ . It then follows that  $C'_0 = C'_1 = 0$ , and hence

$$Z_{\mathcal{X}}^{(\text{base})} = \frac{1}{t^2} (C'_2 t^2 + \dots) \quad \text{as } t \rightarrow 0. \quad (4.41)$$

At such energy scales  $\mu \lesssim R^{-1}$  we thus find

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = C'_2 \log(\Lambda^2/\mu^2) + \text{const.} \quad (4.42)$$

This confirms that for  $\mu \lesssim R^{-1}$  we indeed obtain the expected logarithmic running. *Moreover*, we also see that  $C'_2$  can be identified as the beta-function coefficient for this logarithmic running.

By contrast, let us now consider scales far *above* the KK scale, i.e.,  $\mu \gg R^{-1}$ . In this case we may sum over all of the KK states which are lighter than  $\mu$ . This is the same calculation as in Eq. (3.8), whereupon we find

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = \pi^{\delta/2} R^\delta \int_{\Lambda^{-2}}^{\mu^{-2}} \frac{dt}{t^{1+\delta/2}} Z_{\mathcal{X}}^{(\text{base})}. \quad (4.43)$$

Thus we identify  $C_{j-\delta/2} = \pi^{\delta/2} R^\delta C'_j$ , just as we found for string theory in Eq. (3.34), modulo constants pertaining to the different definitions of the compactification volume. *Moreover*, evaluating our field-theory amplitude, we then find

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = \frac{2}{\delta} \pi^{\delta/2} R^\delta C'_2 (\Lambda^\delta - \mu^\delta), \quad (4.44)$$

as expected. This then reproduces the expected power-law running, with  $C'_2$  now serving as the beta-function coefficient for this running. Indeed,  $C'_2$  plays this role regardless of  $\delta$ —i.e., regardless of the number of compactified extra dimensions which decompactify in the large-volume limit.

This much is field-theoretic. However, for  $\delta > 2$ , we have learned from our theorem [see Eq. (3.25)] that modular invariance and misaligned supersymmetry actually force  $C'_2 = 0$ . Indeed, this result applies for all  $\delta > 2$  (although not  $\delta = 2$  itself). This, then, kills not only the power-law running in Eq. (4.44) in cases with  $\delta > 2$ , but also even the logarithmic running in Eq. (4.42) coming from those sectors. We thus conclude that in any modular-invariant theory which has a decompactification limit in which  $\delta > 2$  extra spacetime dimensions appear, *there is no running at all* from those sectors which are involved in the decompactification process. The contributions from such sectors are therefore completely scale-invariant.

The fact that this running is eliminated ultimately rests on the UV/IR mixing inherent in modular invariance and misaligned supersymmetry. Essentially, for theories with  $\delta > 2$  decompactification limits, our “base” theory (i.e., our theory without the KK/winding excitations) has a nontrivial cancellation purely amongst the zero-mode fields at all mass levels, and it is this cancellation that eliminates the power-law running that would have arisen from the KK excitations of such fields. By the same token, this cancellation also eliminates the running contributions that would have arisen from the winding modes as well.

The case  $\delta = 2$  is somewhat special because we can no longer claim that  $C'_2 = 0$ . Thus, for  $\delta = 2$  we can indeed have logarithmic running for  $\mu < R^{-1}$ . However, as we shall now demonstrate, all power-law running for  $\mu > R^{-1}$  is eliminated even in this case.

To see this, let us begin by noting that for  $\delta = 2$  we have  $C_1 = \pi R^2 C'_2$ . Thus, just as for Eq. (4.44), in field theory we would expect a term of the form

$$\langle \mathcal{X} \rangle_{\text{FT}}^{(4)}(\mu) = \pi R^2 C'_2 (\Lambda^2 - \mu^2) \quad (4.45)$$

where  $\Lambda$  is, as above, a fixed arbitrary cutoff on the Schwinger integral. Unfortunately, such a result is inconsistent with the scale-duality symmetry in Eq. (4.30). Thus, in string theory, the result in Eq. (4.45) must somehow be “completed” to form a scale-duality-invariant quantity. However, as we shall now demonstrate, any such completion will leave us with a fully scale-invariant quantity—i.e., one which has no dependence on  $\mu$  at all. Thus, in this instance, scale duality actually requires scale invariance above the compactification scale.

It is relatively straightforward to construct a scale-duality-invariant “completion” of the expression in Eq. (4.45). Bearing in mind that  $\Lambda$  is an arbitrary but unknown UV cutoff which we expect to naturally be replaced by  $M_s$  within our analysis, we can immediately write a duality-invariant completion of Eq. (4.45) in the form

$$\langle \mathcal{X} \rangle^{(4)}(\mu) = -\pi R^2 C'_2 M_s^2 \times \left[ h(\mu^2/M_s^2) \frac{\mu^2}{M_s^2} + h(M_s^2/\mu^2) \frac{M_s^2}{\mu^2} \right] \quad (4.46)$$

where  $h(x)$  is an unknown function. However, we can immediately list a number of conditions that this  $h$ -function must satisfy the following:

- (i) First, we must demand that the  $\mu \rightarrow 0$  limit of Eq. (4.46) reproduce Eq. (4.45). Specifically, as  $x \rightarrow 0$ , this requires that

$$xh(x) + x^{-1}h(1/x) \approx a_0 + a_1x, \quad (4.47)$$

where  $a_0$  and  $a_1$  are arbitrary constants and where we are identifying  $x \equiv \mu^2/M_s^2$ .

- (ii) Second, because we are restricting our attention to the situation in which only two extra dimensions open up in the decompactification limit, the running of  $\langle \mathcal{X} \rangle^{(4)}(\mu)$  cannot involve terms  $\mu^p$  for any  $p > 2$ . Our claim, then, is that any  $h$ -function satisfying these two conditions must actually have  $a_1 = 0$ . For a function that simply terminates at  $x$ , this indeed follows trivially from the two assumptions above. The fact that the left side of Eq. (4.47) is invariant under  $x \rightarrow 1/x$  in turn implies that  $a_0 + a_1x = a_0 + a_1/x$ , which in turn implies  $a_1 = 0$ .

Of course one may attempt to propose functions for the right side of Eq. (4.47) that behave correctly as  $x \rightarrow 0$  but which would appear to allow a nontrivial dependence on  $x$  in the  $x \rightarrow 0$  limit. For example, one can consider the choice

$$h(x) = \frac{a_0}{x} + a_1 \Theta(1-x) \quad (4.48)$$

which, as  $x \rightarrow 0$ , yields

$$\begin{aligned} & xh(x) + x^{-1}h(1/x) \\ &= a_0 + a_1x\Theta(1-x) + a_1x^{-1}\Theta(1-x^{-1}). \end{aligned} \quad (4.49)$$

If we further assume that the function  $h(x)$  should be differentiable, we can alternatively model the Heaviside  $\Theta$ -function as a sigmoid:

$$\Theta(1-x) \rightarrow \frac{1 + e^{-1}}{1 + e^{x^n-1}} \quad (4.50)$$

for any given  $n > 0$ . We then find that expanding around  $x = 0$  yields  $x^{n+1}$  terms as well as  $x$  terms. However, because we are considering the  $\delta = 2$  case (with only two extra dimensions opening up), we do not expect any physical power-law running beyond quadratic. If we wish to restrict the running to quadratic this in turn restricts us to  $n = 0$ , which renders the expression in Eq. (4.50) equal to a constant.

One can investigate alternative functions but one always runs into similar problems. Essentially, the underlying issue is that the scale-duality symmetry requires that any quadratic running automatically come along with running that has even higher powers. Indeed, if we have any power-law running at all, then the required “turn-over” of this running near the self-dual point  $\mu \approx M_s$  (as required by scale duality) will itself require power-law running involving even higher powers. However, we know from our analysis of the complete amplitude and our result that  $C'_2 = 0$  for  $\delta > 0$  that such higher powers are unphysical. Thus, the theory ensures its own self-consistency by avoiding all power-law running altogether. This argument is, of course, ultimately a consequence of the UV/IR mixing inherent in the scale-duality symmetry. As such, this absence of running—like the scale-duality symmetry itself—is intrinsically a string-theoretic phenomenon.

The end result, then, is that the one-loop running of physical quantities  $\langle \mathcal{X} \rangle^{(4)}(\mu)$  in such theories can at most exhibit the behaviors shown in Fig. 1, where we assume for simplicity that all directions of our compactification tori have equal radii  $R$ . In particular, for  $\delta = 2$  (top panel), the physical quantities in our theory can exhibit at most logarithmic running, and this is limited to the  $\mu \lesssim R^{-1}$  regime. By contrast, for  $\delta > 2$  (bottom panel), these

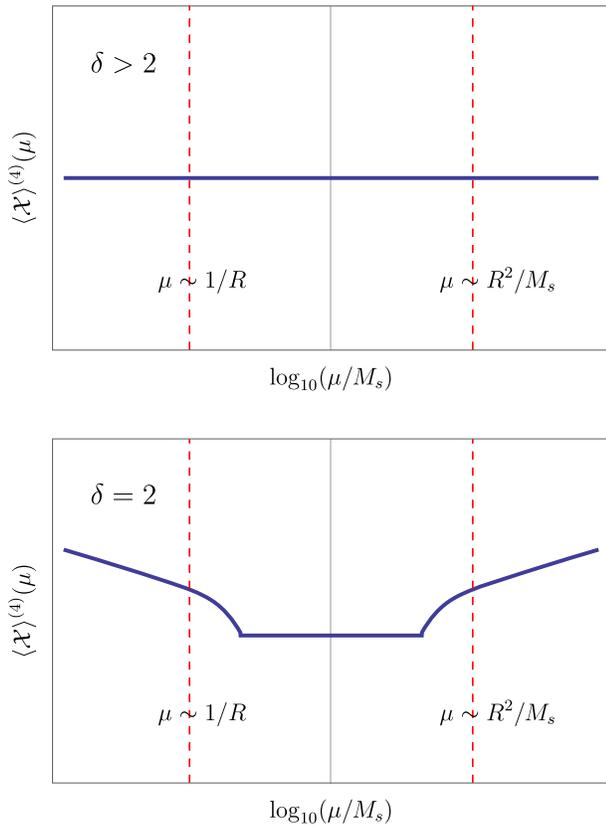


FIG. 1. Sketches of the generic one-loop running behavior for physical quantities that would otherwise run logarithmically in four-dimensional field theory, but now considered within the context of UV/IR-mixed closed string theories in which  $\delta = D - 4$  extra spacetime dimensions open up at a common scale  $R^{-1}$ . For  $\delta = 2$  (top panel), the expected field-theoretic logarithmic running exists only at scales  $\mu \gtrsim R^{-1}$  but (after a string-theoretic transient “pulse” around the scale  $\mu \sim 1/R$ ) is killed beyond this scale and reemerges only for  $\mu \gtrsim M_s^2 R$ , as mandated by the string-theoretic scale-duality symmetry under  $\mu \rightarrow M_s^2/\mu$ . By contrast, for  $\delta > 2$  (bottom panel), the UV/IR mixing is sufficient to kill the running at *all* scales, including the logarithmic running that would have existed at scales  $\mu \lesssim R^{-1}$ . It may seem strange that the emergence of extra spacetime dimensions at a given scale  $\mu_*$  can kill the running at scales below  $\mu_*$ , but this is the direct consequence of the UV/IR mixing and misaligned supersymmetry which connects physics at all scales simultaneously.

quantities cannot even exhibit logarithmic running at any scale, and our quantities remain scale-invariant at all scales.

Thus, to summarize the results of this section, we arrive at the following corollary of our theorem:

*Nonrenormalization corollary.* Within any modular-invariant theory which has  $\delta \equiv D - 4$  large extra dimensions opening up at a scale  $1/R$ , misaligned supersymmetry and UV/IR mixing eliminate all running for  $\mu \gtrsim R^{-1}$  regardless of the value of  $\delta$ . For  $\mu < R^{-1}$ , these same phenomena eliminate all running for  $\delta > 2$ , and leave at most logarithmic running for  $\delta = 2$ .

It is important to understand how this running is eliminated at the level of the actual string spectrum. Precisely what states are cancelling against what other states, as far as their contributions to the overall running are concerned?

The answer depends crucially on the value of  $\delta$ . For all  $\delta > 2$ , as we have seen, the power-law running is cancelled as a result of the vanishing of  $C'_2$ , or equivalently as the result of a vanishing supertrace over the states in the corresponding base theory. This means that there is a cancellation between different species  $A$  and  $B$  of particles in the base theory (or between different *collections*  $A$  and  $B$  of species in the base theory). Of course, viewed from the perspective of our compactification, the base theory contains only the KK and winding zero modes. However, any cancellation between the zero modes of  $A$  and  $B$  particles also naturally extends to the KK/winding excitations associated with these particles as well. Thus, for  $\delta > 2$ , we find that the KK modes associated with species  $A$  cancel against those associated with species  $B$ , while the winding modes associated with species  $A$  cancel against those associated with species  $B$ .

This situation is quite different when  $\delta = 2$ . In this case, we do *not* find that  $C'_2 = 0$ . However, our arguments based on scale-duality invariance tell us that the  $\mu$ -dependence of the power-law running must nevertheless cancel, leaving behind at most a constant. Because  $C'_2 \neq 0$ , this is *not* a cancellation *within* the base theory, but rather a cancellation within its spectrum of KK and winding excitations. In fact, we can be more precise: since the  $A$  and  $B$  zero modes are not cancelling against each other, the KK modes associated with  $A$  are not cancelling with the KK modes associated with  $B$ , nor are the winding modes associated with  $A$  cancelling against the winding modes associated with  $B$ . *Instead*, what we have is a cancellation of the KK modes of  $A$  against the winding modes of  $A$ , and a cancellation of the KK modes of  $B$  against the winding modes of  $B$ .

It may seem strange that a KK contribution can ever cancel against a winding contribution, specially since these contributions will generally have very different scales. However, in a UV/IR-mixed context such as we have here, *all* of our states—both KK and winding—give rise to effects that are simultaneously felt at *all* energy scales in the theory.

We also hasten to add that the cancellations we have been discussing here concern the contributions to the overall *running* of  $\Delta_G(\mu)$ . In particular, these are *not* cancellations of the total contributions to  $\Delta_G(\mu)$  from these states. For example, if the KK states make a contribution of the schematic form  $f(\mu^2/M_s^2)$  to  $\Delta_G(\mu)$  within a certain range of energy scales  $\mu$ , then our assertion is that the corresponding winding states will make a contribution of the schematic form  $C - f(\mu^2/M_s^2)$  to  $\Delta_G(\mu)$  within that same range, where  $C > f(\mu^2/M_s^2)$ . Both contributions have the same overall sign (since the KK and winding states

necessarily have the same spin-statistics), and thus together they produce a total contribution which is nonzero. However, it is the *running* that cancels within this total contribution, leaving  $\Delta_G(\mu)$  entirely  $\mu$ -independent.

This can also be understood at an algebraic level from the results we have outlined above. Recall that our general contributions for  $\delta = 2$  take the form in Eq. (4.45). Within this expression, we may identify the  $\mu^2$  contribution as arising from KK states and the  $\Lambda^2$  contribution (which we may associate with a fundamental high scale such as  $M_s$ ) as arising from winding states. However what the field-theoretic expression in Eq. (4.45) neglects is the fact that this term can itself carry  $\mu$ -dependence. Our argument concerning scale duality then guarantees the cancellation of the  $\mu^2$  contributions coming from these two terms. In other words, a dominant (leading) piece coming from the KK contribution is cancelling against a subdominant piece coming from the winding contribution. This argument is based on the assumption that  $\mu < M_s$ . For  $\mu > M_s$ , by contrast, the roles of KK and winding states are reversed: it is the KK states which give rise to the  $\Lambda^2$  contribution, and it is therefore the subleading KK contributions which cancel against the leading winding contributions.

Thus far, our discussion has focused on factorizable compactifications in which each of the  $\delta$  different compactifications share a common compactification scale  $\mu \sim R^{-1}$ . However, just as with our previous results, these results can be extended to more general situations. These include situations in which the compactification does not have degenerate radii (or more generally compactifications that do not involve a square torus), and also compactifications that are not factorizable.

In the first of these cases the compactification introduces more than one KK scale into the theory. A typical example is a  $\delta$ -dimensional compactification on a  $\delta$ -torus with radii  $R_i$ ,  $i = 1, 2, \dots, \delta$ . Each of these radii is associated with a different KK scale  $M_i \sim R_i^{-1}$ . If the compactification is still factorizable, then our previous discussion still applies. Specifically, at low energies, a state can contribute logarithmic running as usual in four dimensions. However as soon as  $\mu$  reaches the lowest KK scale, all contributions to the running must cancel as embodied in our supertrace identities. Therefore there is no running beyond this point, and indeed all running ceases within the window

$$\min_i(M_i) \ll \mu \ll \max_i(M_s^2/M_i). \quad (4.51)$$

We will see a simple example of this situation in Sec. V, where we consider the case of compactification on a nondegenerate (and nonsquare) two-torus. Of course, for tori with nontrivial shape (complex-structure) moduli as well as nontrivial Kähler moduli, the relevant scales within Eq. (4.51) are simply those associated with the lightest KK

mode and the heaviest winding mode, respectively. In general these can be complicated functions of the radii  $R_i$  due to the nontrivial Kähler moduli.

As indicated above, our primary assertion is that there is no running within the window in Eq. (4.51). That said, there can (and usually will) be string-theoretic transient effects which do not represent true running but which nevertheless lead to localized changes in the values of the relevant amplitudes as we transition *between* different physical regions. This is the “pulse” phenomenon shown in Fig. 1, and we will see explicit examples of such pulses in Sec. V. However, our primary assertion stands: beyond the existence of such pulses, all running ceases above the lightest KK scale.

In the second of these cases, that of nonfactorizable compactifications, we have seen that our theorem applies individually for each of the independent base theories  $Z_s^{(\text{base})}$ . Thus the states in each base theory will either contribute or not contribute to the overall running depending on the physics associated with that sector alone.

We close with a final comment concerning the implications of these results for cases in which the radii  $R^{-1}$  are close to (or even equal to) the string scale. Thus far in this paper, we have been implicitly assuming  $R^{-1} \ll M_s$ —in other words, that we are dealing with large-radius compactifications. However, the results we have obtained are independent of the radii  $R$ , and thus our results will continue to hold even as the KK scales  $R^{-1}$  approach  $M_s$ . This then has some very important implications.

For example, let us consider purely toroidal (untwisted) compactifications of ten-dimensional strings down to four dimensions. The resulting four-dimensional strings clearly have a “decompactification” limit with  $\delta = 6$ . For such an untwisted compactification, this implies that our four-dimensional theory cannot exhibit *any* running—neither logarithmic nor power-law—for  $\mu < R^{-1}$ , or equivalently for the entire range  $0 \leq \mu \leq M_s$ . Of course, if our ten-dimensional theory is supersymmetric, then our resulting four-dimensional theory has  $\mathcal{N} = 4$  supersymmetry. In such a case, our result that this four-dimensional theory exhibits no running is not a surprise and is consistent with existing *supersymmetric* nonrenormalization theorems. However, let us now instead imagine that we are toroidally compactifying the ten-dimensional *non-supersymmetric* tachyon-free  $SO(16) \times SO(16)$  string. In this case, the resulting four-dimensional theory will also be nonsupersymmetric. However, our theorem tells us that this theory will also fail to exhibit any running. Indeed, in this case the UV/IR mixing inherent in misaligned supersymmetry and modular invariance have accomplished what supersymmetry is no longer available to accomplish, namely to suppress the running in the same way as in the supersymmetric case, only now through nontrivial conspiracies across *all* of the states in the string spectrum rather than through boson/fermion pairings.

Of course, this result applies only for untwisted compactifications. For twisted compactifications, the situation described in Sec. III F implies that certain contributions to the running will also be eliminated. These are the contributions from the sectors that are involved in decompactification limits with  $\delta > 2$ . However, according to our theorem, logarithmic running may yet emerge from the sectors with  $\delta = 2$ . This feature is also reminiscent of what occurs in supersymmetric contexts. However, once again, we are finding that this also holds in *non*-supersymmetric contexts, purely as a result of UV/IR mixing.

We consider this observation to be quite general. It is well known that various supersymmetric theories are governed by nonrenormalization theorems. What we are finding, however, is that analogous nonsupersymmetric theories *also* seem to be governed by the same sorts of nonrenormalization theorems. Thus UV/IR mixing appears to play many of the same roles as supersymmetry.

## V. EXPLICIT EXAMPLE: GAUGE COUPLINGS IN $\mathbb{T}^2$ STRING COMPACTIFICATIONS

The discussion in the previous sections was completely general, and is thus applicable to all tachyon-free closed string theories. However, it is instructive to revisit a well-known example and perform an explicit calculation within the framework of that example in order to see how the results of our theorem are realized in practice, and how the cancellation of power-law running actually occurs. Along these lines, this will also allow us to understand the critical role played by UV/IR mixing, and in particular by our insistence on maintaining worldsheet modular invariance throughout our calculation.

### A. Modular-invariant calculation versus traditional calculation

For this purpose, we shall focus on a calculation which was historically at the center of an early triumph in string phenomenology, namely the first calculations of gauge threshold contributions within an  $\mathcal{N} = 1$  supersymmetric heterotic string model which is compactified from ten dimensions to four dimensions on an orbifolded six-dimensional torus. This calculation was first performed in Ref. [37] using the formalism established in Ref. [38]. The results of this analysis were subsequently developed in many later works and proved highly influential. However, as we shall discuss, the analysis in Refs. [37,38] did not respect the full (worldsheet) modular invariance of the theory. As a result, this calculation did not capture all of the features which we are analyzing in this paper. By redoing this calculation within our fully modular-invariant framework, our goal is to see how the new features we have been discussing arise. This includes the running of the gauge couplings, and the absence thereof beyond the appropriate

KK scales. For this reason, this calculation will prove to be a useful testing ground for our work.

As in Ref. [37], we shall consider the special case in which the six-dimensional torus can be factorized into a four-dimensional torus  $\mathbb{T}^4$  and a two-dimensional torus  $\mathbb{T}^2$ . With certain further assumptions outlined in Ref. [37], we may disregard the physics associated with the  $\mathbb{T}^4$  compactification. This problem then reduces to a study of the two-dimensional compactification of an  $\mathcal{N} = 1$  six-dimensional closed heterotic string to four dimensions on an orbifolded two-torus  $\mathbb{T}^2$ .

Finally, since our goal will be to compare our results with those of Ref. [37], we stress that we will actually be calculating a quantity which is distinct from that calculated in Ref. [37]. It will therefore be important to understand the difference between these quantities when attempting to make comparisons.

In general, our goal is to calculate  $\Delta_G$ , i.e., the one-loop contribution to  $16\pi^2/g_G^2$ . Here  $G$  is any of the unbroken gauge-group factors within the string model in question,  $g_G$  is the corresponding gauge coupling, and  $\Delta_G$  is defined in Eq. (2.17). Moreover, we recall that prior to the introduction of a regulator we can identify  $\Delta_G = \langle \mathcal{X} \rangle$  where  $\mathcal{X}$  is given in Eq. (2.8) with the operator insertions  $\mathbb{X}_\ell$  given in Eq. (2.18). The quantity  $\Delta_G$  will be finite if there are no massless  $\mathbb{X}_2$ -charged states in the string spectrum [i.e., assuming that Eq. (3.1) continues to hold].

For this calculation, however, such states generically *do* appear. Thus Eq. (3.1) will no longer be valid, and a regulator will be needed. As discussed in Sec. IV B, we shall adopt the modular-invariant regulator function  $\mathcal{G}_\rho(a, \tau)$  described in Sec. II B. We therefore seek to calculate the finite quantity

$$\Delta_G \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} Z_{\mathcal{X}} \mathcal{G}_\rho(a, \tau) \quad (5.1)$$

where  $\mathcal{X} = \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2$  [with the  $\mathbb{X}_\ell$  defined in Eq. (2.18)]. Moreover, through the identification in Eq. (4.20), we may also view this as a scale-dependent quantity  $\Delta_G(\mu)$ .

In Ref. [37], by contrast, a somewhat different quantity  $\Delta_G^{(\text{DKL})}$  is calculated. For each gauge group  $G$ , the quantity  $\Delta_G^{(\text{DKL})}$  represents a *threshold correction* for the running of the corresponding gauge coupling  $16\pi^2/g_G^2$ —a correction which arises due to the infinite towers of massive states in the string spectrum. As such,  $\Delta_G^{(\text{DKL})}$  tallies the contributions to the running of  $16\pi^2/g_G^2$  from only the *massive* states in the theory. As explained in Ref. [38], this quantity is defined as

$$\Delta_G^{(\text{DKL})} \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} (Z_{\mathcal{X}'} - \tau_2 b_G) \quad (5.2)$$

where  $\mathcal{X}' \equiv \tau_2^2 \mathbb{X}_2$  and where  $b_G$  is given in Eq. (4.26). The subtraction of  $\tau_2 b_G$  within the integrand of Eq. (5.2) reflects the idea that  $\Delta_G^{(\text{DKL})}$  includes the contributions of only the massive states in the theory. However, as we have described above, this subtraction also has the further benefit of rendering  $\Delta_G^{(\text{DKL})}$  finite. Indeed, given the form of  $\mathbb{X}_2$ , we see that  $b_G$  is nothing but the one-loop beta-function coefficient for  $G$ .

We thus see that there are three important differences between our expression for  $\Delta_G$  and the corresponding expression  $\Delta_G^{(\text{DKL})}$  that is calculated in Ref. [37]. First, we see that the operator insertion within  $\Delta_G^{(\text{DKL})}$  is truncated, including only the  $\mathbb{X}_2$  term but disregarding the  $\mathbb{X}_1$  term. Second, we see that these expressions utilize different methods of ensuring a finite result: our expression in Eq. (5.1) utilizes a smooth, modular-invariant regulator function  $G_\rho(a, \tau)$  that suppresses the potentially divergent contributions from the large- $\tau_2$  region of integration, while the expression in Eq. (5.2) utilizes a sharp, brute-force subtraction of the otherwise divergent contribution from the massive  $\mathbb{X}_2$ -charged states. Finally,  $\Delta_G^{(\text{DKL})}$  has no energy scale associated with it, but is simply a threshold that must be matched to an effective theory. By contrast, our  $\Delta_G$  is calculated within the complete theory, with the regulator itself defining the energy scale.

As a result of these differences,  $\Delta_G$  is modular invariant while  $\Delta_G^{(\text{DKL})}$  is not. Indeed,  $\Delta_G$  includes the operator insertion  $\mathbb{X}_1$ , and  $\mathbb{X}_1$  is the modular completion of  $\mathbb{X}_2$ , as discussed in Ref. [17]. Likewise, the regulator function  $\mathcal{G}_\rho(a, \tau)$  within  $\Delta_G$  eliminates the divergences that would have arisen if  $b_G \neq 0$ , but does so in a fully modular-invariant way. By contrast, the brute-force subtraction of massless contributions within  $\Delta_G^{(\text{DKL})}$  breaks modular invariance, since the UV/IR mixing within modular invariance generally causes the massless states to mix non-trivially with all of the other states in the theory, thereby rendering such a targeted subtraction impossible.

At first glance, it may seem surprising that  $\Delta_G^{(\text{DKL})}$  is not modular invariant. After all, the well-known expression for this quantity that is ultimately derived in Ref. [37] (and which we shall also obtain below) turns out to be a modular function of the compactification moduli. However, as described above, we are calculating these quantities within the framework of a six-dimensional theory which is toroidally compactified to four dimensions. There are therefore *two* distinct modular symmetries that we expect to play a role in this calculation—not only the usual *worldsheet* modular invariance that must be unbroken for all perturbative closed string theories, but also the *spacetime* modular invariance associated with such a two-dimensional  $\mathbb{T}^2$  toroidal compactification. Indeed, while  $\Delta_G^{(\text{DKL})}$  respects the spacetime modular invariance, it lacks the worldsheet modular invariance that is our main interest

in this paper. As such, it does not respect the stringy UV/IR mixing that drives our theorem and its consequences. By contrast,  $\Delta_G$  preserves both modular symmetries simultaneously, as it must. It is therefore only within the calculation of the properly defined  $\Delta_G$  that we expect the consequences of our theorem to become manifest.

## B. General setup

We begin our evaluation of Eq. (5.1) by recalling the physical setup. As discussed above, we assume a six-dimensional heterotic string with  $\mathcal{N} = 1$  spacetime supersymmetry compactified on a two-torus  $\mathbb{T}^2$ . Our goal is to calculate the amplitude  $\Delta_G \equiv \langle \mathcal{X}' \rangle^{(4)}$ . To do this, we shall work within the approximation that our compactification volume is large. This will not affect our final results but will allow us to make contact with the calculation in Ref. [37]. Within this approximation, we can then utilize the result for  $\langle \mathcal{X}' \rangle^{(4)}$  given in Eq. (4.29) for  $\delta = 2$ , specifically

$$\langle \mathcal{X}' \rangle^{(4)} \approx \frac{\pi}{3} \tilde{V}_T (C'_2 - b_G) + b_G \int_{\mathcal{F}} \frac{d^2 \tau}{\tau_2^2} \tau_2 Z_{\text{KK}/\text{winding}} \mathcal{G}_\rho(a, \tau), \quad (5.3)$$

where in the present case we have from Eqs. (2.18) and (A9) the coefficient

$$C'_2 = -2 \text{Str}'(Q_G^2 \bar{Q}_H^2) + \frac{1}{6} \text{Str}'_E Q_G^2 - \frac{\xi}{2\pi} \text{Str}'(\bar{Q}_H^2 \tilde{M}^2) + \frac{\xi}{24\pi} \text{Str}'_E \tilde{M}^2. \quad (5.4)$$

Before proceeding further, several comments are in order. First, we note that the terms in the second line of Eq. (5.4) correspond to the so-called “ $\mathcal{Y}$ -terms” in Refs. [37,38]. Historically these terms tended to be disregarded because they are universal (i.e., independent of the gauge group in any given string model) and therefore play no role in gauge coupling unification (which is sensitive to only the *differences* of gauge couplings). Moreover, these terms receive contributions from only the massive string states in the base theory and are thus insensitive to any potential running of the gauge couplings at scales below  $M_s$ . However, as we have seen, these terms are part of the full modular-invariant calculation and we shall therefore retain these terms within  $C'_2$ .

Second, we also note that the contributions from the massless states within Eq. (5.4) are nothing but the corresponding beta-function coefficient:

$$C'_2|_{M=0} = b_G \equiv -2 \text{Str}'_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2. \quad (5.5)$$

Indeed, these terms are nothing other than the divergent contribution  $\text{Str}'_{M=0} \mathbb{X}_2$ .

Of course, for our purposes our main interest in this section is in the second line of Eq. (5.3) because this is the term that pertains to  $\Delta_G^{(\text{DKL})}$ . This is therefore the term which will allow us to compare our results with those of Ref. [37]. Moreover, the first line of Eq. (5.3) has no dependence on the regulator and therefore will ultimately have no  $\mu$ -dependence. This is therefore a  $\mu$ -independent constant. We shall therefore henceforth focus on the expression

$$I_G \equiv b_G \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2} \tau_2 Z_{\text{KK/winding}} \mathcal{G}_\rho(a, \tau), \quad (5.6)$$

so that the total one-loop contribution  $\Delta_G$  to the inverse coupling  $g_G$  is given by

$$\Delta_G = \langle \mathcal{X} \rangle^{(4)} = \frac{\pi}{3} \tilde{V}_T (C'_2 - b_G) + I_G. \quad (5.7)$$

Our goal, then, is to evaluate  $I_G$  for compactifications on a two-torus. To do this, we must know the physical spectrum of our theory. We will therefore follow standard conventions by taking the general moduli for the two-torus to be written as  $T = T_1 + iT_2$  and  $U = U_1 + iU_2$ , where for reference the metric and  $B$ -field are defined in terms of the moduli as

$$G_{ij} = \frac{T_2}{U_2} \begin{pmatrix} 1 & U_1 \\ U_1 & |U|^2 \end{pmatrix}, \quad B_{ij} = T_1 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (5.8)$$

Note that for  $T_1 = U_1 = 0$  (corresponding to rectangular tori without a Kalb-Ramond  $B$ -field), we can identify

$$T_2 = M_s^2 R_1 R_2 \quad \text{and} \quad U_2 = \frac{R_2}{R_1}. \quad (5.9)$$

Although it will sometimes be useful for pedagogical purposes to think in terms of such rectangular tori with radii  $R_1$  and  $R_2$ , in the following we shall nevertheless let all four moduli be nonzero. However, without loss of generality, we shall adopt the convention that  $U_2 > 1$ .

Given this, the  $\delta = 2$  KK/winding factor  $Z_{\text{KK/winding}}$  in Eq. (5.6) is simply given by  $Z_{\mathbb{T}^2}$ , where

$$Z_{\mathbb{T}^2}(\tau) = \sum_{\vec{k}, \vec{\ell} \in \mathbb{Z}^2} e^{-\pi\tau_2 \alpha' M_{\mathbb{T}^2}^2} e^{2\pi i \tau (k_2 \ell_1 - k_1 \ell_2)} \quad (5.10)$$

and where the squared-mass contribution coming from KK/winding states with KK/winding numbers  $k_1, k_2, \ell_1, \ell_2 \in \mathbb{Z}$  is given by

$$\alpha' M_{\mathbb{T}^2}^2 = \frac{|k_1 + Uk_2 + T\ell_1 + TU\ell_2|^2}{U_2 T_2}. \quad (5.11)$$

Before we present the fully modular-invariant expression for the gauge-coupling contribution, it is useful to recall for

comparison the result of Ref. [37] which was derived by computing the integral in Eq. (5.2). This gives

$$\begin{aligned} \Delta_G^{(\text{DKL})} &= b_G \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2} (\tau_2 Z_{\mathbb{T}^2}(\tau) - \tau_2) \\ &= -b_G \log \left( \frac{8\pi e^{1-\gamma_E}}{3\sqrt{3}} T_2 U_2 |\eta(T)|^4 |\eta(U)|^4 \right), \end{aligned} \quad (5.12)$$

where  $b_G$  is given by the coefficient in Eq. (5.5) and where  $\gamma_E$  is the Euler-Mascheroni constant. We emphasize again that this  $\tau_2$  subtraction, which is done to regulate the integral by removing the logarithmic divergence coming from the massless states, breaks worldsheet modular invariance. Therefore this result does not reflect the worldsheet modular invariance of the full theory from which it came. However, as this  $\tau_2$ -subtraction is independent of  $T$  and  $U$ , this integral correctly reflects the *spacetime* modular symmetry associated with the spacetime toroidal compactification. This spacetime modular symmetry is manifest in the  $U$ - and  $T$ -dependence of the one-loop gauge-coupling correction, and is inherited from the worldsheet modular symmetry.

Our objective, of course, is to obtain an expression for the equivalent term in  $\Delta_G$  (namely  $I_G$ ) which preserves not only spacetime modular invariance but also worldsheet modular invariance. It is to this task that we now turn.

### C. Summary of results

Before plunging into our analysis, it may help to provide a bird's-eye view by summarizing our eventual results.

Our main result, of course, is the evaluation of the full modular-invariant integral in Eq. (5.6). We shall find that this is highly nontrivial but eventually yields the expression

$$\begin{aligned} I_G &= \frac{-b_G}{1+a^2\rho} \left\{ \log(cT_2 U_2 |\eta(T)\eta(U)|^4) + 2\log\sqrt{\rho}a \right. \\ &\quad \left. + \frac{8}{\rho-1} \sum_{\gamma, \gamma' \in \Gamma_\infty \setminus \Gamma} \left[ \tilde{\mathcal{K}}_0^{(0,1)} \left( \frac{2\pi}{a\sqrt{(\gamma \cdot T_2)(\gamma' \cdot U_2)}} \right) \right. \right. \\ &\quad \left. \left. - \frac{1}{\rho} \tilde{\mathcal{K}}_1^{(1,2)} \left( \frac{2\pi}{a\sqrt{(\gamma \cdot T_2)(\gamma' \cdot U_2)}} \right) \right] \right\} \end{aligned} \quad (5.13)$$

where we have adopted the shorthand notation  $\gamma \cdot T_2 \equiv \text{Im}(\gamma \cdot T)$ , where we have defined

$$c \equiv 16\pi^2 \rho^{-\frac{\rho+1}{\rho-1}} e^{-2(\gamma_E+1)}, \quad (5.14)$$

and where we have defined the Bessel-function combinations [13]

$$\tilde{\mathcal{K}}_\nu^{(n,p)}(z, \rho) \equiv \sum_{k,r=1}^{\infty} (krz)^n \left[ K_\nu(krz/\rho) - \rho^p K_\nu(krz) \right]. \quad (5.15)$$

Here  $K_\nu(z)$  are modified Bessel functions of the second kind, and  $\gamma$  and  $\gamma'$  are the *spacetime* modular

transformations acting in the respective spacetime modular groups associated with  $T$  and  $U$ . Likewise,  $\Gamma_\infty \setminus \Gamma$  is the set of modular transformations which collectively “unfold” the fundamental domain of the full modular group  $\Gamma$  into the “strip” with  $|\tau_2| \leq 1/2$  and  $\tau_1 > 0$ . Indeed, this strip is nothing but the fundamental domain of the modular subgroup  $\Gamma_\infty$  which is generated only by the transformation  $\tau \rightarrow \tau + 1$  and which therefore leaves the cusp at  $\tau = i\infty$  invariant.

As we see, the first line of the expression for  $I_G$  in Eq. (5.13) clearly contains the classic moduli-dependent pieces of Eq. (5.12). However, we now see that this result also comes with the additional terms on the second and third lines of Eq. (5.13). These extra terms not only maintain the spacetime modular invariance that already existed in the first line, but also restore worldsheet modular invariance for the entire amplitude, as required.

It is not difficult to recover the result for  $\Delta_G^{(\text{DKL})}$  in Eq. (5.12) from the full expression for  $I_G$  in Eq. (5.13). To see this, we note that within the result of Eq. (5.13) the quantities  $(\rho, a)$  parametrize the regulator function  $G_\rho(a, \tau)$  which appears in Eq. (5.6). Therefore  $\Delta_G^{(\text{DKL})}$  is equivalent to  $I_G$  with the logarithmic piece  $\log \sqrt{\rho a}$  subtracted by hand and with  $a$  sent to zero. In this limit the Bessel functions all vanish, and the remaining difference between the threshold of Eq. (5.12) and the  $a \rightarrow 0$  limit of Eq. (5.13) with subtracted  $\log \sqrt{\rho a}$  amounts to a difference in renormalization scheme. This scheme-dependence is encapsulated within the parameter  $c$ . Indeed, by numerically equating the remaining terms we find that equivalent schemes would correspond to  $\rho \approx 22$ .

Given the result in Eq. (5.13), and following the procedures outlined at the beginning of Sec. IV B, we can also analyze how  $I_G$  and thus the full  $\Delta_G$  runs in the modular invariant theory. Identifying the physical mass scale  $\mu$  as in Eq. (4.20), we find that the running behavior for  $I_G$  can most easily be described by partitioning the full range of  $\mu$  into five different regimes. For presentational purposes, we shall assume that our  $\delta = 2$  compactification geometry consists of a rectangular two-torus with radii  $R_1$  and  $R_2$  with  $R_2 \gg R_1$  and with no Kalb-Ramond field. We then find that our five separate regions of interest are given by

$$\begin{aligned}
 \text{Region I: } & \mu \ll 1/R_2, \\
 \text{Region II: } & 1/R_2 \ll \mu \ll 1/R_1, \\
 \text{Region III: } & 1/R_1 \ll \mu \ll M_s^2 R_1, \\
 \text{Region IV: } & M_s^2 R_1 \ll \mu \ll M_s^2 R_2, \\
 \text{Region V: } & \mu \gg M_s^2 R_2.
 \end{aligned} \tag{5.16}$$

Within each of these regions and far from the boundaries *between* these regimes, we can then evaluate the

approximate leading behaviors for  $I_G$ . As we shall see, this ultimately yields the results

Region I:

$$I_G \approx \frac{\pi}{3} \left( M_s^2 R_1 R_2 + \frac{R_2}{R_1} \right) - 2 \log(\mu R_2),$$

Region II:

$$I_G \approx \frac{\pi}{3} \left( M_s^2 R_1 R_2 + \frac{R_2}{R_1} \right),$$

Region III:

$$I_G \approx \frac{\pi}{3} M_s^2 R_1 R_2,$$

Region IV:

$$I_G \approx \frac{\pi}{3} \left( M_s^2 R_1 R_2 + \frac{R_2}{R_1} \right),$$

Region V:

$$I_G \approx \frac{\pi}{3} \left( M_s^2 R_1 R_2 + \frac{R_2}{R_1} \right) - 2 \log \left( \frac{M_s^2 R_2}{\mu} \right). \tag{5.17}$$

These results provide confirmation of many of our previous assertions. For example, within Regions II through IV, we see that there is no logarithmic or power-law running at all. By contrast, within Region I we have at most a logarithmic running which essentially ceases as we cross from Region I to Region II and encounter the lightest KK states. Moreover, as expected, we observe that Regions IV and V are directly related to Regions II and I respectively under the scale-duality transformation in Eq. (4.30), while Region III is self-dual. Thus the absence of running that we observed in Regions II and III extends into Region IV, with logarithmic running reappearing only in Region V.

Two further comments are in order. First, while the above results indicate the leading behaviors for our overall amplitudes  $I_G(\mu)$ , the results for Regions II and IV actually contain an overall coefficient which is weakly  $\mu$ -dependent. This is the overall coefficient in Eq. (4.17), or equivalently the coefficient  $1/(1+a^2\rho) = 1/(1+\mu^2/M_s^2)$  in Eq. (5.13). However, this dependence is not a true running, but simply represents a residual  $\mu$ -dependence which is subleading within these regions. By contrast, for Region III, this residual  $\mu$ -dependence is exactly cancelled within our result for Region III. Thus, within Region III, our result is fully  $\mu$ -independent, exhibiting no running at all and remaining truly constant.

Second, we observe that we have limited our analysis to the behaviors of  $I_G(\mu)$  within the *interiors* of each region in Eq. (5.16). However, interesting behavior can also be found at the boundaries between these different regions. In particular, we shall find that between Regions II and III, and also between Regions III and IV, we have a transient

“pulse.” This is not a pure “running,” but rather a completely stringy phenomenon which serves to connect the constant value of  $I_G$  which emerges on one side of each of these boundaries with the different constant value of  $\Delta_G$  which emerges on the other.

For the record, we emphasize that the above behavior is fairly generic for  $\delta = 2$ , which has the potential for two different compactification radii  $R_1$  and  $R_2$ . Of course, for  $\delta > 2$ , we have the potential for additional distinct regions opening up due to the possible appearance of additional KK scales. Indeed, the behavior within all of these regions would continue to be arranged in a manner consistent with the scale-duality symmetry. However, as we have explained in Sec. IV B, we would find that  $I_G$  is a constant in each region.

This concludes the summary of our main results. Much of the rest of this section is devoted to explaining how these results can be extracted from our general result in Eq. (5.13). However, it turns out that this extraction will entail a number of interesting subtleties and technical maneuvers, some of which revolve around writing our general result in Eq. (5.13) in a series of different but equivalent ways. For this reason, the reader who does not wish to delve into those details and who is willing to accept the above results can skip directly to Sec. V G.

#### D. Evaluating the one-loop contribution to the modular-invariant gauge coupling

Let us start by explicitly deriving the fully modular-invariant result of Eq. (5.13). Our derivation proceeds by evaluating the integral in Eq. (5.6), where for convenience we will henceforth set  $b_G = 1$ , since  $b_G$  merely provides an overall factor.

In order to perform this calculation we first need to fully specify the physical spectrum over which we will eventually be taking supertraces. In general, the states appearing in Eq. (5.11) are not level-matched. Therefore we must first determine the spectrum of physical states. Given the expression in Eq. (5.10), these are the states which satisfy the level-matching constraint

$$k_2 \ell_1 - k_1 \ell_2 = 0. \quad (5.18)$$

This condition straightforwardly yields two sets of solutions for the physical spectrum:

$$(k_1, k_2, \ell_1, \ell_2) \in A \cup B \quad (5.19)$$

where

$$\begin{aligned} A &\equiv \left\{ (k_1, k_2, 0, 0) \mid (k_1, k_2) \in \mathbb{Z}^2 \right\}, \\ B &\equiv \left\{ (c\tilde{k}_1, c\tilde{k}_2, d\tilde{k}_1, d\tilde{k}_2) \mid (\tilde{k}_1, \tilde{k}_2) \in \mathbb{Z}^2, \right. \\ &\quad \left. \gcd(\tilde{k}_1, \tilde{k}_2) = 1, (c, d) \in \mathbb{Z}^2, d \geq 1 \right\}. \end{aligned} \quad (5.20)$$

As described in Sec. IV B, knowledge of this spectrum then allows us to determine the fully modular-invariant term  $I_G(\mu)$ . As shown in Eq. (4.18), this contribution to the amplitude can naturally be expressed in terms of a “reduced” amplitude  $P(a)$ , which from Eq. (4.19) now takes the form

$$P(a) = \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \tau_2 Z_{\mathbb{T}^2}(\tau) Z_{\text{circ}}(a, \tau) \quad (5.21)$$

where  $Z_{\text{circ}}$  is defined in Eq. (3.5). Following the methods outlined in Ref. [13], we shall evaluate  $P(a)$  by writing

$$P(a) = P'_1(a) + P'_2(a) \quad (5.22)$$

where

$$\begin{aligned} P'_1(a) &= \frac{1}{a} \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \tau_2 Z_{\mathbb{T}^2}(\tau) - \frac{1}{a} \int_1^\infty \frac{d\tau_2}{\tau_2} \\ P'_2(a) &= P_2(a) + \frac{1}{a} \int_1^\infty \frac{d\tau_2}{\tau_2} \end{aligned} \quad (5.23)$$

with

$$P_2(a) = \frac{2}{a} \int_0^\infty \frac{d\tau_2}{\tau_2^2} g_{\mathcal{X}}(\tau_2) \sum_{\ell=1}^\infty e^{-\pi\ell^2/(a^2\tau_2)}. \quad (5.24)$$

Note that in Eq. (5.23), the final terms for  $P'_1(a)$  and  $P'_2(a)$  cancel in the sum  $P(a)$ . However, the reason for introducing these terms is that they allow us to shuffle logarithmic divergences between  $P'_1(a)$  and  $P'_2(a)$ , with the arbitrary finite parameter  $t$  allowing these two quantities to be independently convergent. It will be an important self-consistency check on our calculation that all dependence on  $t$  will naturally drop out of the sum  $P(a)$ . As discussed in Ref. [13], this reshuffling does not have to preserve modular invariance in the individual terms because modular invariance is ultimately restored in the sum.

Note that we can identify  $P'_1(a)$  as the “traditional” non-modular-invariant minimally subtracted integral in Eq. (5.12). The remaining quantity  $P'_2(a)$  thus contributes all the “extra” terms in Eq. (5.13) which render our effective cutoffs smooth and moreover restores worldsheet modular invariance to our calculation.

As discussed in Ref. [13], the integral for  $P'_1$  can be performed using the Rankin-Selberg-Zagier techniques [20–22] which, upon adapting the results of Ref. [39] and performing the sum over the physical spectrum in Eq. (5.20), gives

$$P'_1(a) = -\frac{1}{a} \log(4\pi T_2 U_2 |\eta(T)\eta(U)|^4) + \frac{1}{a} \log(e^{\gamma_E t}). \quad (5.25)$$

To complete the calculation for  $P(a)$ —and indeed to restore the worldsheet modular invariance—we must add the  $P'_2(a)$  integral to this result. Following Ref. [13] and utilizing the same “unfolding” techniques as above, this integral can also be expressed as a sum over the physical spectrum in Eqs. (5.11) and (5.20). This then yields a total expression for  $P(a)$ :

$$\begin{aligned}
 P(a) = & -\frac{1}{a} \log\left(4\pi T_2 U_2 |\eta(T)\eta(U)|^4\right) \\
 & + \frac{8}{a} \sum_{k,r=1}^{\infty} \sum_{\gamma, \gamma' \in \Gamma_{\infty} \setminus \Gamma} K_0\left(\frac{2\pi}{a} \frac{rk}{\sqrt{(\gamma \cdot T_2)(\gamma' \cdot U_2)}}\right) \\
 & - \frac{1}{a} \log(4\pi a^2 e^{-2\gamma_E})
 \end{aligned} \quad (5.26)$$

where we have adopted the same conventions and notations as described below Eq. (5.15). As required, this result is now invariant under worldsheet modular symmetries. It is also finite for every  $a > 0$ ; it is independent of  $t$ , as required; and it also exhibits spacetime modular invariance for the compactification moduli  $T$  and  $U$ .

Note that for practical purposes the sums over  $r, k$  can be combined with the sums over the coset to yield the result

$$\begin{aligned}
 P(a) = & -\frac{1}{a} \log(4\pi T_2 U_2 |\eta(T)\eta(U)|^4) \\
 & + \frac{8}{a} \sum_{c,d,c',d'} K_0\left(\frac{2\pi}{a} \frac{|cT + d||c'U + d'|}{\sqrt{T_2 U_2}}\right) \\
 & - \frac{1}{a} \log(4\pi a^2 e^{-2\gamma_E})
 \end{aligned} \quad (5.27)$$

where  $(c, d) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$  and  $(c', d') \in \mathbb{Z}^2 \setminus \{(0, 0)\}$ .

Having determined the reduced amplitude  $P(a)$ , all that remains is to insert this into Eq. (4.18), whereupon the full  $I_G$  given in Eq. (5.13) is obtained.

### E. Running gauge couplings

Thus far we have calculated the one-loop contribution to the gauge coupling as a fixed quantity, much as was originally done in Ref. [37] except that we have done this in a manner that fully respects worldsheet modular invariance. However, as we have discussed above and in Refs. [13, 17], we may now go one step further and proceed to interpret our regulator variables  $\rho$  and  $a$  as defining a running scale  $\mu$  through the identification  $\mu \equiv \sqrt{\rho} a M_s$ . In this way we can then interpret our results as yielding a *running* coupling in the low-energy effective field theory derived from the string. To map out the behavior of this running, in this section we will analyze the result in Eq. (5.13) in various limits and in various energy windows.

In order to derive the behavior of  $I_G$  in various limits and regions, as in Eq. (5.17), it will prove convenient to work with the reduced amplitude  $P(a)$  of Eq. (5.26) rather than

the full  $I_G$  of Eq. (5.13), as the former is far simpler to analyze. However it is important to appreciate that the identification made in Eq. (4.20) of the energy scale  $\mu$  with  $\sqrt{\rho} a M_s$  is a crucial physical step in going from  $P(a)$  to  $\Delta_G(\mu)$ , as discussed at length in Ref. [13]. Indeed due to the scale-duality symmetry in Eq. (4.30), the alternative identification  $\mu \equiv M_s/(\sqrt{\rho} a)$  would be an equally valid choice. Making this choice is tantamount to choosing which direction of our worldsheet theory should be identified as UV versus IR physics in spacetime. As discussed in Ref. [13], it is inevitable that such a choice must be made in order to extract an EFT from our underlying UV/IR-mixed string theory.

We shall now proceed to evaluate  $P(a)$ —and ultimately  $I_G(\mu)$ —within the different regions outlined in Eq. (5.16). To do this, in each case we shall find an approximation for  $P(a)$  that is valid within the appropriate region, and thereby deduce the leading running that emerges within that region. This requires some care, as the approximations that are appropriate in each case are very different from each other. We shall therefore be relatively explicit in how these approximations are made in each case.

#### 1. Region I: Field-theory limit

Having chosen to identify  $\mu = \sqrt{\rho} a M_s$ , let us now as a first step examine the low-energy behavior of Region I which extends to  $\mu \rightarrow 0$ , or equivalently to  $\sqrt{\rho} a \ll 1$ . Here we expect that Eq. (5.13) yields the effective four-dimensional field-theoretic behavior together with finite gauge thresholds. Indeed, we know that in this limit the regulator should “turn off” entirely, i.e.,  $\mathcal{G}_\rho(a, \tau) \rightarrow 1$ , whereupon the original unregulated integral should be restored. Of course the unregulated integral had a logarithmic divergence that emerges because  $b_G \equiv \text{Str} \mathbb{X}_2 \Big|_{M=0} \neq 0$ . Thus we expect that in this limit Eq. (5.13) gives the threshold correction result in Eq. (5.12) plus a term that diverges logarithmically as  $\mu \rightarrow 0$ .

This is indeed what happens. Recognizing the Bessel-function asymptotic behavior  $K_\nu(z) \sim \sqrt{\pi/(2z)} e^{-z}$  as  $z \rightarrow \infty$ , we see that the terms involving Bessel functions do not survive in this limit. We then find by direct analysis of Eq. (5.13) that the only remaining terms are

$$I_G(\mu) \approx -\log(cT_2 U_2 |\eta(T)\eta(U)|^4) - 2 \log\left(\frac{\mu}{M_s}\right) \quad (5.28)$$

as  $\mu \rightarrow 0$ . Assuming  $T_2 \gg 1$  and recalling that  $\eta(iz) \sim e^{-\pi z/12}$  for  $z \gg 1$ , we can extract the leading volume dependence of  $I_G(\mu)$ . Up to additional terms of order unity, we obtain

$$I_G(\mu) \approx \frac{\pi}{3} (T_2 + U_2) - \log\left(\frac{\mu^2}{M_s^2} T_2 U_2\right) \quad (5.29)$$

as  $\mu \rightarrow 0$ . This approximation is in accord with the  $T$ -volume scaling relation in Eq. (5.6), and is valid in the region  $a \ll 1/\sqrt{T_2 U_2}$  or equivalently  $\mu \ll 1/R_2$  for the rectangular torus, yielding the Region I behavior described in Eq. (5.17). Note that this is the energy scale below which Kaluza-Klein modes are effectively inactive and may be considered to have been integrated out.

It is important to understand the differences between our modular-invariant result  $I_G(\mu)$  in Eq. (5.28) and the traditional result  $\Delta^{(\text{DKL})}$  of Eq. (5.12). Indeed, explicitly restoring the factors of  $b_G$ , we see that the two are related for  $\mu \ll M_s$  via

$$I_G(\mu) = -b_G \log \frac{\mu^2}{M_s^2} - b_G \log \left( \frac{3\sqrt{3}c}{8\pi e^{1-\gamma_E}} \right) + \Delta^{(\text{DKL})}. \quad (5.30)$$

Thus, while  $I_G(\mu)$  is a contribution to the amplitude that is a function of a running mass scale  $\mu$ , we see that  $\Delta^{(\text{DKL})}$  is a scale-independent *threshold correction*. Of course, as we have repeatedly stressed,  $\Delta_G^{(\text{DKL})}$  is not a modular-invariant quantity. It is only the full expression  $I_G(\mu)$  which respects the full modular invariance of the theory.

We thus see that our modular-invariant calculation not only keeps track of the running, but also keeps track of the natural degrees of freedom that are dynamical in the EFT associated with the scale  $\mu$ . As explained in more detail in Refs. [13] and [17], for any scale  $\mu$  our regulator implicitly keeps track of which states can be classified as either “light” (with masses  $M \lesssim \mu$ ) or “heavy” (with masses  $M \gtrsim \mu$ ): the contributions from the heavy states are suppressed by our regulator and have thus effectively been integrated out, while the contributions from the light states are retained. Indeed, this establishes  $\mu$  as a floating mass scale, which in turn enables us to take  $\mu$  to  $M_s$  and beyond.

## 2. Region V: Ultrahigh energies

Let us now consider what happens in the  $\mu \rightarrow \infty$  limit. In order to do this we note that, as discussed in Ref. [13], the scale-duality symmetry  $\mu \rightarrow M_s^2/\mu$  requires that the reduced amplitude  $P(a)$  have an  $a \rightarrow 1/a$  symmetry, and this is in turn ensured by the explicit  $a \rightarrow 1/a$  symmetry of  $Z_{\text{circ}}(a)$ . Thus in order to study the  $a \rightarrow \infty$  limit we can simply replace  $a \rightarrow 1/a$  in Eq. (5.26), yielding a dual  $P(a)$  of the form

$$P(a) = -a \log(4\pi T_2 U_2 |\eta(T)\eta(U)|^4) - a \log \left( \frac{4\pi e^{-2\gamma_E}}{a^2} \right) \quad (5.31)$$

up to Bessel-function terms which are exponentially suppressed as  $a \rightarrow \infty$ . This then yields

$$I_G \approx -\log(cT_2 U_2 |\eta(T)\eta(U)|^4) + 2 \log \left( \frac{\mu}{M_s} \right) \quad (5.32)$$

as  $\mu \rightarrow \infty$ . Thus for  $T_2 \gg 1$  the leading volume dependence up to additional terms of order unity becomes

$$I_G(\mu) \approx \frac{\pi}{3}(T_2 + U_2) + \log \left( \frac{\mu^2}{M_s^2} T_2 U_2 \right) \quad \text{as } \mu \rightarrow \infty. \quad (5.33)$$

Indeed, for the rectangular torus this approximation is valid for  $a \gtrsim \sqrt{T_2 U_2}$ , or equivalently for

$$\mu \gtrsim R_2 M_s^2, \quad (5.34)$$

yielding the Region V behavior of Eq. (5.17).

## 3. Regions II, III, and IV: Stringy energies

Since both the  $\mu \rightarrow 0$  and  $\mu \rightarrow \infty$  limits display only logarithmic behavior, one might then suspect that power-law running behavior is present near the self-dual point  $a = 1$ . Indeed, motivated by the field-theoretic result in Eq. (4.45), one may even suspect that in this region the contribution from  $P'_2$  cancels that from  $P'_1$  so as to give  $I_G \rightarrow 0$  at the self-dual point.

To investigate this, let us focus more closely on the  $a \approx 1$  region by writing  $P'_2$  in the alternative form

$$\begin{aligned} P'_2(a) = & \frac{1}{a} \log(4\pi T_2 U_2 |\eta(T)|^4 |\eta(U)|^4) - \frac{1}{a} \log(e^{\gamma_E} t) \\ & - \sqrt{T_2 U_2} \log \left| \eta \left( ia \sqrt{T_2/U_2} \right) \eta \left( ia^{-1} \sqrt{T_2/U_2} \right) \right|^4 \\ & - 4 \sum_{\substack{\gamma \in \Gamma_\infty \\ \gamma \neq 1}} \sum_{k=1}^{\infty} \left\{ \sqrt{T_2(\gamma \cdot U_2)} \log \left( 1 - e^{-2\pi a k \sqrt{\frac{T_2}{\gamma \cdot U_2}}} \right) \right. \\ & \left. + \sqrt{U_2(\gamma \cdot T_2)} \log \left( 1 - e^{-2\pi a k \sqrt{\frac{U_2}{\gamma \cdot T_2}}} \right) \right\} \\ & + \frac{8}{a} \sum_{\substack{\gamma \in \Gamma_\infty \\ \gamma \neq 1}} \sum_{k,r=1}^{\infty} K_0 \left( \frac{2\pi}{a} \frac{rk}{\sqrt{(\gamma \cdot T_2)(\gamma' \cdot U_2)}} \right). \end{aligned} \quad (5.35)$$

Note that in writing Eq. (5.35) we have neglected terms that vanish in the derivative definition of  $I_G(\mu)$  in Eq. (4.18). Rendering the sums explicitly, we then find that this becomes

$$\begin{aligned} P'_2(a) = & \frac{1}{a} \log(4\pi T_2 U_2 |\eta(T)|^4 |\eta(U)|^4) - \frac{1}{a} \log(e^{\gamma_E} t) \\ & - \sqrt{T_2 U_2} \log \left| \eta \left( ia \sqrt{T_2/U_2} \right) \eta \left( ia^{-1} \sqrt{T_2/U_2} \right) \right|^4 \\ & - 4 \sum_{\substack{c,d \\ c \neq 0}} \left\{ \frac{\sqrt{T_2 U_2}}{|cU+d|} \log \left( 1 - e^{-2\pi a |cU+d| \sqrt{T_2/U_2}} \right) \right. \\ & \left. + \frac{\sqrt{T_2 U_2}}{|cT+d|} \log \left( 1 - e^{-2\pi a |cT+d| \sqrt{U_2/T_2}} \right) \right\} \\ & + \frac{8}{a} \sum_{\substack{c,d,c',d' \\ c,c' \neq 0}} K_0 \left( \frac{2\pi |cT+d| |c'U+d'|}{a \sqrt{T_2 U_2}} \right), \end{aligned} \quad (5.36)$$

where as before  $c, d, c', d' \in \mathbb{Z}$ .

Before proceeding further, we observe that the calculation that leads to this result requires considerable care. In particular, we must regulate all of the terms using methods similar to those in Ref. [37]. The steps in this calculation are as follows. First one splits the spectrum of  $M_{\mathbb{T}^2}^2$  in Eq. (5.11) into three sets: the first has  $l_1 = l_2 = 0$  and coincides with the  $A$  set in Eq. (5.20); the second has  $k_2 = l_2 = 0$ ; and the third has  $l_2 \neq 0$ . In performing the calculation for the first two sets we Poisson-resum  $(k_1, k_2)$  and  $(l_1, k_1)$  respectively. Then, performing a partial sum over the greatest common divisors of  $(k_1, k_2)$  and  $(l_1, k_1)$  and subtracting the double-counted contribution from the states with  $l_1 = l_2 = k_2 = 0$  gives the first four lines of Eq. (5.35). Finally, the contribution from the third set of states is computed directly without any Poisson resummation, and yields the Bessel functions on the final line. These simply resemble the equivalent terms appearing in the previous incarnation of  $P'_2$  in Eq. (5.26).

The expression in Eq. (5.36) is most useful for considering  $P'_2$  near  $\mu \approx M_s$ . Moreover, the first line of Eq. (5.36) precisely cancels  $P'_1$  in Eq. (5.25). Thus, at first sight it seems that the one-loop contribution to the gauge-coupling could indeed cancel in this region in a manner that mimics Eq. (4.45).

However, this is not the case. To see this, let us consider the window

$$\frac{1}{\sqrt{T_2 U_2}} \ll a \ll \sqrt{T_2 U_2}. \quad (5.37)$$

For rectangular tori this corresponds to

$$1/R_2 \ll \mu \ll M_s^2 R_2, \quad (5.38)$$

which may be taken as defining an “extended stringy regime” that encapsulates Regions II, III, and IV. By inspection we find that within this window the terms in the third through fifth lines of Eq. (5.35) are all exponentially suppressed. Therefore, neglecting these exponentially suppressed terms and adding  $P'_2$  to  $P'_1$  in Eq. (5.25) we obtain a very compact approximation for  $P(a)$  in this region:

$$P(a) \approx -\sqrt{T_2 U_2} \log \left| \eta \left( ia \sqrt{\frac{T_2}{U_2}} \right) \eta \left( ia^{-1} \sqrt{\frac{T_2}{U_2}} \right) \right|^4. \quad (5.39)$$

Defining for convenience a dimensionless energy scale  $\hat{\mu} \equiv \mu/M_s$ , we find that Eq. (4.18) near  $\hat{\mu} \approx 1$  gives

$$I_G \approx \frac{\pi}{3} T_2 \frac{E_2 \left( \frac{i}{\hat{\mu}} \sqrt{\frac{T_2}{\rho U_2}} \right)}{(1-\rho)(1+\hat{\mu}^2)} + \rho \rightarrow \frac{1}{\rho} + \hat{\mu} \rightarrow \frac{1}{\hat{\mu}} + (\rho, \hat{\mu}) \rightarrow \left( \frac{1}{\rho}, \frac{1}{\hat{\mu}} \right) \quad (5.40)$$

where  $E_2$  is the Eisenstein function defined in Eq. (A3) and where each substitution within the right side of Eq. (5.40) is meant to operate purely on the first term (thereby leaving us with a total of only four terms on the right side of this equation). Given the definition of the Eisenstein function, we see that for large  $T_2$  all the  $E_2$  factors are unity up to terms that are exponentially suppressed as long as  $M_s \sqrt{U_2/T_2} \ll \mu \ll M_s \sqrt{T_2/U_2}$ . Of course, this is nothing but Region III. Thus, within Region III we find that Eq. (5.40) is exponentially well approximated by

$$I_G(\mu) \approx \frac{\pi}{3} T_2 \quad \text{for } \mu \approx M_s. \quad (5.41)$$

Recalling that  $T_2$  is the  $T$ -volume  $\tilde{V}_T$ , we thus verify, as anticipated, that there is no quadratic  $\mu$ -dependence—or indeed running of any kind—within this energy-scale window. Indeed, within this region,  $I_G(\mu)$  is exponentially well approximated by a constant.

Finally, going beyond this region, we see that Eqs. (5.29), (5.33), and (5.41) together imply that  $I_G(a) \approx (\pi/3)T_2$  everywhere. Indeed, this holds regardless of the energy scale  $\mu$  so long as  $U_2 \approx 1$ .

The full analytical expression within Eq. (5.40) can also be evaluated numerically. In Fig. 2 we plot an example of the resulting running behavior for  $\Delta_G(\mu)$  that emerges in the case of a square torus with  $T_2 = 100$  and  $U_2 = 1$  within the extended stringy regime of Eq. (5.38) (blue curve). The fact that this torus is square implies that our two KK scales  $1/R_1$  and  $1/R_2$  collapse to become the same scale, thereby rendering Regions II and IV nonexistent. We also plot the corresponding beta function (red curve). We observe, as promised, that there is essentially no running

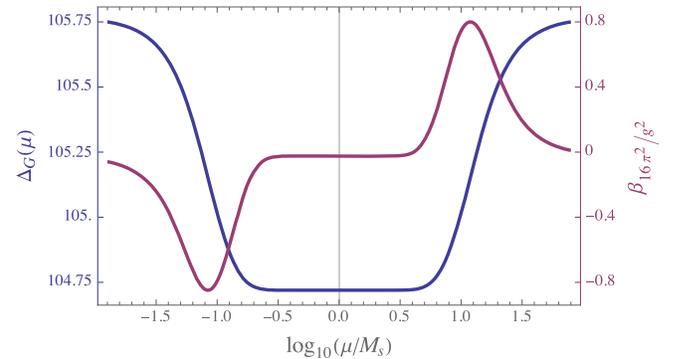


FIG. 2. The precise one-loop contribution to  $16\pi^2/g^2$  (blue) in Eq. (5.40), along with the corresponding beta function (red) in Eq. (4.21), plotted as functions of the energy scale  $\mu$  within the extended stringy regime of Eq. (5.38) for a toroidal  $\mathbb{T}^2$  compactification with compactification moduli  $T_2 = 100$  and  $U_2 = 1$ . These curves are evaluated using the full expression in Eq. (5.39), and provide an explicit example of the behavior anticipated in Fig. 1. Note that we are actually plotting  $I_G(\mu)$  rather than  $\Delta_G(\mu)$ , but these quantities differ by a scale-independent constant given by Eq. (5.7) which is irrelevant for our purposes.

within this regime; for example, within the range  $|\log(\mu/M_s)| \lesssim 0.7$  our blue curve is completely flat. However, for  $|\log(\mu/M_s)| \gtrsim 0.7$  we see that there exists a small transient ‘‘pulse’’ as the theory crosses the KK thresholds on either side of this regime, in accord with the behavior anticipated in Fig. 1. We also observe that the string-scale value of  $I_G$  differs from the low-energy value near  $\mu \approx 0$  by a term that grows with volume only as  $\log T_2$  rather than as the total number of Kaluza-Klein modes with masses below the string scale. Indeed, the latter behavior would have been expected if there had been a dominant region of power-law running between the low scale and  $M_s$ .

This behavior changes if our compactification torus is not square. As an example, let us consider the case in which our compactification torus is *rectangular*, with  $U_2 \gg 1$  but a purely imaginary shape modulus. In this case Regions II and IV open up and become part of our extended stringy regime, potentially inducing an arbitrarily large splitting of the KK scales associated with each of the two dimensions of the torus. In such cases, Regions II, III, and IV become separated within the extended stringy regime. Within this extended regime there is suppressed running of  $I_G(\mu)$  with transient behavior localized near the boundaries. Once again, field-theoretic running ceases as soon as the first KK states are encountered, as described in Sec. IV.

We can learn more about the nature of this transient ‘‘pulse’’ by concentrating on Regions II and IV. Recall that our expressions for  $P(a)$  and  $I_G(\mu)$  in Eqs. (5.39) and (5.40) respectively were valid across the entire extended stringy regime in Eq. (5.38) consisting of Regions II, III, and IV. By contrast, the approximate result in Eq. (5.41) was valid only within Region III. Indeed, within Region II the argument appearing in  $\eta(ia\sqrt{T_2/U_2})$  is no longer large. However, within this window we may modular-transform the first Dedekind eta function of Eq. (5.39) to find yet another form for  $P(a)$  given by:

$$P(a) \approx -\sqrt{T_2 U_2} \log \left| \eta \left( ia^{-1} \sqrt{\frac{U_2}{T_2}} \right) \eta \left( ia^{-1} \sqrt{\frac{T_2}{U_2}} \right) \right|^4 - \sqrt{T_2 U_2} \log \left( a^2 \frac{U_2}{T_2} \right). \quad (5.42)$$

Neglecting constant terms that vanish under the derivative definition of  $\Delta_G$  in Eq. (4.18), this then becomes

$$P(a) \approx \frac{\pi}{3} \frac{1}{a} (T_2 + U_2) - 2\sqrt{T_2 U_2} \log a. \quad (5.43)$$

Given the  $\log a$  term within Eq. (5.43), it may at first sight seem that we have discovered linear power-law running within Region II. Indeed the individual terms in Eq. (4.18) are proportional to  $a^2 \partial_a P(a) = a$ , which, when multiplied by the prefactor  $\sqrt{T_2 U_2} \equiv M_s R_2$ , would imply linear contributions to  $I_G$  that are proportional to  $\mu R_2$ .

As we have discussed above, this would appear to make sense because Region II is an approximately five-dimensional regime, with  $1/R_2 \ll \mu \ll 1/R_1$ . However, what actually appears in the definition of  $I_G$  in Eq. (4.18) is not a single  $P(a)$  but rather the combination  $P(\rho a) - P(a)$ , and within this difference these logarithm terms cancel. Indeed, simple  $\log(a)$  terms within  $P(a)$  do not make contributions to  $I_G$ . This is not a coincidence, since the appearance of the combination  $P(\rho a) - P(a)$  is ultimately dictated by the scale-duality symmetry. We therefore conclude that we do not have any running in Region II. Needless to say, the same conclusion also applies to the dual Region IV.

## F. Putting it all together

Pulling together all of our previous results, we may summarize our expressions for  $P(a)$  within each of the five regions corresponding to Eq. (5.16) as follows:

Region I:  $a \ll 1/\sqrt{T_2 U_2}$

$$P(a) \approx \frac{\pi}{3} \frac{1}{a} (T_2 + U_2) - \frac{1}{a} \log(16\pi^2 e^{-2\gamma_E} a^2 T_2 U_2),$$

Region II:  $1/\sqrt{T_2 U_2} \ll a \ll 1/\sqrt{T_2/U_2}$

$$P(a) \approx \frac{\pi}{3} \frac{1}{a} (T_2 + U_2),$$

Region III:  $1/\sqrt{T_2/U_2} \ll a \ll \sqrt{T_2/U_2}$

$$P(a) \approx \frac{\pi}{3} \left( a + \frac{1}{a} \right) T_2,$$

Region IV:  $\sqrt{T_2/U_2} \ll a \ll \sqrt{T_2 U_2}$

$$P(a) \approx \frac{\pi}{3} a (T_2 + U_2),$$

Region V:  $a \gg \sqrt{T_2 U_2}$

$$P(a) \approx \frac{\pi}{3} a (T_2 + U_2) - a \log(16\pi^2 e^{-2\gamma_E} a^{-2} T_2 U_2). \quad (5.44)$$

Note that within these expressions we have omitted terms which do not yield any contributions to the corresponding  $I_G$ . These include not only constant terms but also terms scaling as  $\log a$ , as discussed above.

These expressions generate the suite of behaviors that were summarised in Eq. (5.17). Moreover, given our expressions for  $I_G(\mu)$ , we find that the same running applies to  $\Delta_G(\mu)$  since these two quantities differ only by a scale-independent constant, as shown in Eq. (5.7). Thus, proceeding upwards in energy scale, we find the following:

- (i) At the lowest energy scales (Region I), we find that  $\Delta_G(\mu)$  evolves as  $\log \mu$ . This behavior assumes that our underlying theory has  $\text{Str } \mathbb{X}_2 \neq 0$ ; otherwise this logarithmic running is absent.

- (ii) Next, after crossing the lowest KK threshold (Region II), we find that  $\Delta_G(\mu)$  flattens out and becomes  $\mu$ -independent.
- (iii) Next, as we approach the second KK threshold,  $\Delta_G(\mu)$  experiences a “pulse” and then enters Region III. Within Region III,  $\Delta_G(\mu)$  is again flat, but with a different constant value. In general, this pulse is a genuinely transient effect that appears when we flow *between* (rather than *within*) our asymptotic regions. The size of this pulse depends on the shape modulus associated with this toroidal compactification; for a square (or rectangular) torus this magnitude is approximately  $\pi U_2/3$ .
- (iv) The above behavior persists all the way to the string scale  $\mu = M_s$ . Beyond this scale, the behavior of  $\Delta_G(\mu)$  is fixed by scale-inversion duality, with  $\Delta_G(\mu) = \Delta_G(M_s^2/\mu)$ .

This running is shown in Fig. 3, along with two other runnings: the minimally subtracted EFT result in which the four-dimensional coupling runs purely logarithmically

across all energy scales and is simply matched to  $\Delta_G^{(\text{DKL})}$  at the scale  $\mu = M_s$ , and the traditional field-theoretic power-law running that is expected from the accumulated effects of the KK modes beyond the first KK threshold at  $\mu = R_2^{-1}$ . For this example we have chosen values of  $T_2 = 10^4$  and  $U_2 = 100$  (on a rectangular torus) in order to make the different behaviors in Regions I, II, and III all evident.

As evident from Fig. 3, all of these different runnings share a common feature, specifically a logarithmic running at extremely low energies  $\mu \ll M_s$  within Region I. However, once we cross the first KK threshold at  $\mu = 1/R_2$ , the behaviors diverge. The four-dimensional EFT approach is of course completely insensitive to these thresholds that correspond to heavy modes because this approach is tantamount to a simple matching procedure. As such, it is incapable of capturing any physics at energy scales that exceed the lowest KK scale. Meanwhile the traditional approach which involves summing the contribution of the KK states experiences a strong power-law

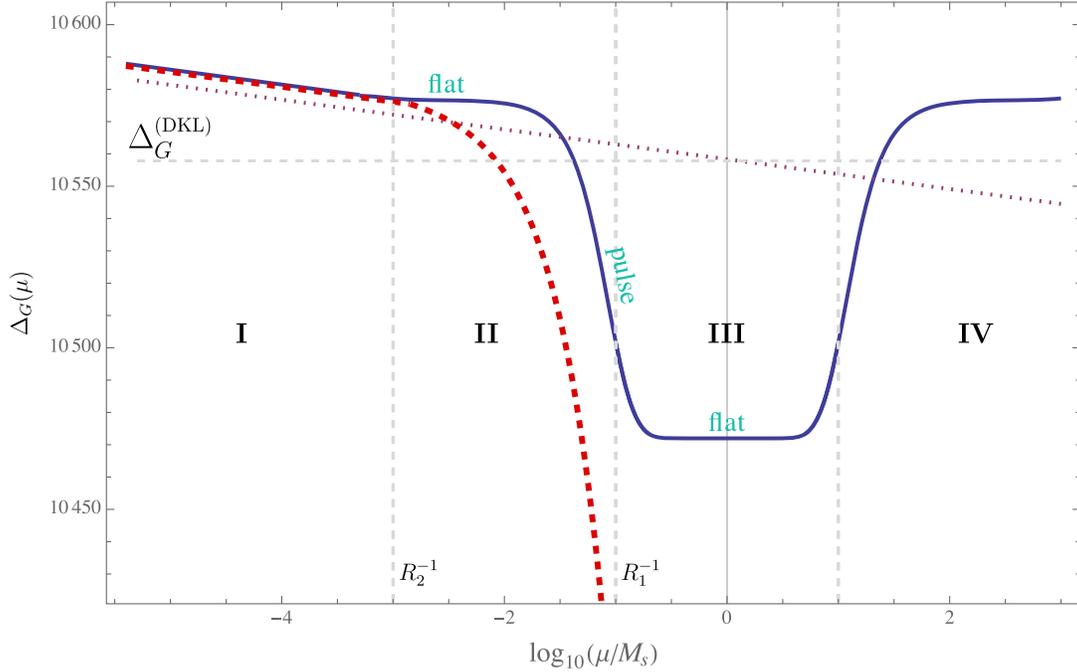


FIG. 3. Comparison of the results of three different running calculations. The first (faint dotted red line) is the minimally subtracted EFT result in which the four-dimensional coupling runs purely logarithmically across all energy scales and is simply matched to  $\Delta_G^{(\text{DKL})}$  at the scale  $\mu = M_s$ . The second (dotted red curve) indicates the traditional field-theoretic power-law running that is expected from the accumulated effects of the KK modes beyond the first KK threshold at  $\mu = R_2^{-1}$ , with our theory becoming effectively five-dimensional in Region II and eventually six-dimensional in Region III. By contrast, the third (solid blue curve) shows what actually emerges when the full worldsheet modular invariance of our theory is taken into account, including the use of a fully modular-invariant regulator (in this case,  $\mathcal{G}_\rho$  with  $\rho = 2$ ). As indicated, all running ceases beyond the first KK threshold, except for the appearance of a transient “pulse” localized between Regions II and III. All three plots correspond to choosing a rectangular compactification torus with  $T_2 = 10000$  and  $U_2 = 100$ . Despite the fact that all three runnings exhibit logarithmic behavior at low scales  $\mu \ll 1/R_2$ , we see that they diverge significantly at higher energies, with the fully modular-invariant result ultimately leading to a fixed-point regime at  $\mu \approx M_s$  and a dual regime beyond  $M_s$ . This result also automatically avoids the Landau pole that might otherwise appear below  $M_s$ . Once again, as in Fig. 2, we are actually plotting  $I_G(\mu)$  rather than  $\Delta_G(\mu)$ , but these quantities differ by a scale-independent constant in Eq. (5.7) which is irrelevant for our purposes.

running above the lowest KK scale. Indeed, in the example shown in Fig. 3, this running even leads to a Landau pole below the string scale. The existence of such a Landau pole would signal a fundamental inconsistency, indicating that our theory becomes nonperturbative at such energy scales. If this were true, a one-loop perturbative approach would no longer be consistent.

Fortunately, our fully modular-invariant calculation leads to an entirely different behavior. Indeed, we see from Fig. 3 that all running ceases above  $\mu = 1/R_2$ . As a result,  $\Delta_G(\mu)$  becomes flat. In Fig. 3 this flat behavior is evident within the lower portion of Region II. However, we see from Fig. 3 that this flat behavior soon gives way to our transient “pulse” near the second threshold at  $\mu = 1/R_1$ . This pulse is an entirely “stringy” phenomenon which smoothly connects the flat behavior in Region II to the flat behavior in Region III, causing a sudden drop in the value of  $\Delta_G(\mu)$ . Despite appearances, we emphasize that this drop in the value of  $\Delta_G(\mu)$  is relatively small (a mere 1% effect) in the example shown. Moreover, this pulse is only a transient effect in the sense that its duration does not grow into the asymptotic regions, but instead is localized to the boundary between Regions II and III. Finally, upon entering Region III, the running then remains flat all the way to the string scale  $M_s$ . Beyond this, the running is governed by the entrance into a dual phase of the theory beyond  $M_s$ .

The properties shown in Fig. 3 have two important implications. First, the existence of the scale duality between the  $\mu < M_s$  and  $\mu > M_s$  regions indicates that there is a fundamental limit on the degree to which our theory can exhibit such field-theoretic UV behavior. This is extensively discussed in Refs. [13,17], and will be discussed further in the Conclusions.

Second, the cessation of running beyond the lightest KK scale serves to protect our theory from the possible appearance of Landau poles. In fact, the scale-duality symmetry even potentially keeps our couplings perturbative *over the entire range* from the deep IR up to  $M_s$  and beyond. This therefore gives our perturbative calculations an intrinsic self-consistency over the entire range of  $\mu$ .

### G. Interplay of KK and winding states

We conclude this section with one final comment concerning the roles of Kaluza-Klein states and winding states in our analysis.

As we have seen, there is no logarithmic or power-law running for  $\Delta_G(\mu)$  within Regions II, III, and IV. Indeed, within Region III, it is an exact result that  $\Delta_G(\mu)$  contains no  $\mu$ -dependence at all. Moreover, as we explained below Fig. 1 (at the very end of Sec. IV B), the cessation of running in this  $\delta = 2$  case arises because the contributions from Kaluza-Klein states are ultimately cancelling against those of winding states. While Kaluza-Klein states begin to appear as we cross from Region I to Region II, at first glance it might appear that winding states would not appear

until the crossing from Region III into Region IV. However, in a UV/IR-mixed theory, this conclusion is too narrow—indeed, the effects of both Kaluza-Klein and winding states are felt *throughout* the string spectrum. Indeed, it is precisely because of this mixing that the cancellation of running occurs at all, and that it persists throughout the entire extended stringy regime stretching from Region II through Region IV.

Given our previous results, this can also be understood at an algebraic level. Indeed, the two factors within the logarithm that appears in Eq. (5.39) can be identified with Kaluza-Klein and winding modes respectively. Accordingly, within Region III, the KK modes make a contribution to the gauge coupling correction of

$$\Delta_G(\mu)\Big|_{\text{KK}} = \frac{\pi}{3} \frac{\mu^2/M_s^2}{1 + \mu^2/M_s^2} R_1 R_2. \quad (5.45)$$

This quantity is proportional to the number of “active” KK modes at the scale  $\mu$  (i.e., the number of KK modes with masses less than  $\mu$ ), as one would expect from power-law running in a two-dimensional compactification. However, the corresponding winding-mode contribution is given by

$$\Delta_G(\mu)\Big|_{\text{winding}} = \frac{\pi}{3} \frac{1}{1 + \mu^2/M_s^2} R_1 R_2. \quad (5.46)$$

We thus see that the overall  $\mu$ -dependence cancels in the sum, leading to the  $\mu$ -independent quantity quoted for Region III in Eq. (5.17).

Interestingly, for  $\mu < M_s$ , the power-law running contribution to  $\Delta_G$  in Eq. (5.45) does not dominate; rather, it is the corresponding winding-mode contribution in Eq. (5.46) which dominates. By contrast, for  $\mu > M_s$ , these roles are exchanged: the Kaluza-Klein modes dominate, and it is the winding modes that contribute the subdominant piece.

## VI. CONCLUSIONS, DISCUSSION, AND FUTURE DIRECTIONS

In this paper we have derived a new renormalization theorem for theories that exhibit UV/IR mixing. The natural setting for our theorem, namely theories exhibiting modular invariance, includes all perturbative closed string theories. It has been known for a long time that in four dimensions all such tachyon-free theories satisfy a remarkable supertrace constraint [2]

$$\text{Str } \mathbf{1} = 0 \quad (6.1)$$

where the supertrace—as suitable for theories with *infinite* towers of states—has the regulated definition given in Eq. (2.33). Of course, if the theory in question exhibits spacetime supersymmetry, this identity is satisfied level-by-level, with the contribution from each state cancelling in pairwise fashion against that of its superpartner. However,

even if the theory in question lacks spacetime supersymmetry, Eq. (6.1) continues to hold exactly as a consequence of a hidden *misaligned supersymmetry* [1–3] that exists within the spectrum of all tachyon-free modular-invariant string theories. This misaligned supersymmetry is the manifestation of the underlying UV/IR mixing inherent in modular invariance. Moreover, in the absence of spacetime supersymmetry, the constraint in Eq. (6.1) is not satisfied through pairwise cancellations; indeed, no such pairwise cancellations are possible. Instead, in such cases Eq. (6.1) holds as the result of a nontrivial cancellation amongst the contributions from *all* of the states across the *entire* spectrum of the theory—exactly as one would expect in a UV/IR-mixed theory.

In complete analogy with the constraint in Eq. (6.1), another similar “core” constraint relates the one-loop cosmological constant (or more precisely, the zero-point one-loop amplitude)  $\Lambda$  of any tachyon-free modular-invariant theory to a supertrace over its physical mass spectrum [2]:

$$\Lambda = \frac{1}{24} \mathcal{M}^2 \text{Str} M^2, \quad (6.2)$$

where  $\mathcal{M}$  is the reduced string scale  $M_s/(2\pi)$ . Indeed, this constraint holds for all of the theories in which Eq. (6.1) holds, with both constraints ultimately emerging together within the same analysis [2].

As we have seen, the constraints in Eqs. (6.1) and (6.2) have important physical ramifications. It turns out that the first of these constraints can actually be regarded as a constraint on the zero-point one-loop amplitude  $\Lambda$ . Indeed, within a purely field-theoretic context, this identity tells us that the one-loop quartic divergence that would ordinarily have arisen for the one-loop cosmological constant is “magically” cancelled—indeed, cancelled *exactly*. However, within a string-theory context, Eq. (6.1) is actually sufficient to kill *all* of the divergences of  $\Lambda$ , not only the quartic divergence that might have been expected in field theory. Because of this, the value of  $\Lambda$  in string theory is actually *finite*, and this finite value is then given in Eq. (6.2).

This cancellation and the results in Eqs. (6.1) and (6.2) arise regardless of the particular phenomenology that the string might exhibit (such as its gauge symmetry or particle content) [2]. These constraints are even preserved in the face of radiative corrections or if the theory passes through phase transitions in which the fundamental degrees of freedom change, so long as modular invariance is preserved. Indeed, these results require only the preservation of modular invariance and the absence of physical tachyons, or equivalently the existence a misaligned supersymmetry in the string spectrum. As such, we regard these constraints as a core part of the UV-completeness of the theory.

At first glance, it might appear that these constraints hold only at one-loop order. However, this guess would not be

correct. To understand this, we must first realize that modular invariance itself is an *all-orders* symmetry. This point might seem somewhat counter-intuitive, since modular invariance is *motivated* by the requirement that one-loop amplitudes be consistent with worldsheet reparametrization invariance. Indeed, modular invariance is the symmetry that ensures that reparametrization invariance continue to hold for genus-one diagrams—and not only genus-zero diagrams—despite the extra “large” transformations that are possible around the noncontractible cycles of the torus. For this reason this symmetry is sometimes called “one-loop” modular invariance. However, the important point is that this symmetry—regardless of its motivations—is an exact symmetry which must be enforced exactly within any string theory. Indeed, if it were not enforced, one-loop diagrams would not be consistent. Of course, the requirement that the two-loop amplitudes also be consistent with reparametrization invariance might provide *additional* constraints on our theory. However, such constraints would merely *augment* the one-loop constraints. They would not replace the constraints from modular invariance any more than the constraints of modular invariance replace those of reparametrization invariance at tree level.

Given this, one might wonder what masses appear within the supertrace in Eq. (6.2). The answer is relatively simple. To any order in perturbation theory, our theory will have a spectrum of (potentially radiatively corrected) masses and a (potentially radiatively corrected) gravitational background. Within this background, there will likewise be a corresponding (potentially corrected) value of the zero-point genus-one amplitude  $\Lambda$ . Our claim, then, is that these new masses and this new value for  $\Lambda$  will continue to be exactly related by Eq. (6.2), and that Eq. (6.1) will continue to hold as well. Indeed, this is why we regard these two constraints as fundamental universal truths: their validity holds regardless of the order at which we are performing our calculations, and stems directly from modular invariance in any tachyon-free theory.

In this paper, we have presented a new theorem which likewise governs the spectra of a certain class of four-dimensional tachyon-free modular-invariant theories—those that have decompactification limits. Like the above constraints, this theorem is also completely general, holding regardless of the particle content and gauge symmetries of the underlying theory, and like the above constraints it also ultimately arises because of the nontrivial cancellations inherent in misaligned supersymmetry. Moreover, as we have demonstrated, this theorem also yields many additional spectral constraints which are cousins of those in Eqs. (6.1) and (6.2).

This theorem holds for modular-invariant tachyon-free theories in which there is at least one decompactification limit, i.e., a limit in which a geometric compactification volume can be taken to infinity, resulting in a higher-dimensional theory. Such four-dimensional theories can

therefore be viewed as geometric compactifications of higher-dimensional theories. As we have described in Sec. III A, our theorem rests on the observation that there is a subtle mathematical clash that arises within the modular structure of such theories as the volume of compactification of the four-dimensional theory is taken to infinity. Resolving this clash implies that the original four-dimensional theory must satisfy not only the universal constraint in Eqs. (6.1)—a constraint which is appropriate for four dimensions—but also certain additional constraints that are appropriate for the higher-dimensional theory. These additional constraints further restrict the properties of the theory, and likewise affect many more amplitudes than just the cosmological constant. Moreover, these additional conditions are independent of the radius of compactification.

Even more interestingly, we found that these additional constraints immediately lead to a new *nonrenormalization* theorem for our four-dimensional theory. This theorem also applies for all compactification radii, and holds even without spacetime supersymmetry. Like its antecedents, this theorem also holds for large classes of physical quantities, including (but not limited to) the case of gauge couplings.

The implications of this nonrenormalization theorem are different for each possible decompactification limit experienced by our four-dimensional theory. This is ultimately the case because (as we have discussed in Sec. IV B) each decompactification limit contributes its own constraints that must hold in the original four-dimensional theory.

For limits involving decompactifications in which  $\delta > 2$  new dimensions open up, our theorem implies that there is *no running at all* from the states involved in the sectors producing that limit. In other words, the states which survive the limit do not collectively yield any running at all, regardless of the compactification radius associated with that limit.

By contrast, for limits involving decompactifications in which only  $\delta = 2$  new dimensions open up, the implications of our theorem are more nuanced. In this case, the states from such sectors can give rise to at most a logarithmic running, but this can exist only at scales below the lightest of the KK thresholds associated with that limit. Beyond this critical scale, we once again find that there are no collective contributions to the running.

For a given string model involving combinations of decompactification limits, these results generally imply that there is never any running at all beyond the lightest of the relevant KK scales. Indeed, the only changes in the value of the relevant amplitude beyond the lightest of the KK scales are those which arise as the result of our string-theoretic “pulse.” This effect is transient and highly localized to certain KK thresholds. As such, this pulse is not a true running.

Along with our nonrenormalization theorem, in this paper we have also proven two “ $T$ -volume scaling” rules

which apply when the compactified volume is large but *not* infinite. The first of these rules applies to finite amplitudes and asserts that the amplitudes in this regime are always proportional to the product of the decompactified  $(4 + \delta)$ -dimensional amplitude and a so-called “ $T$ -duality-invariant compactification volume”  $\tilde{V}_T$ . This result is given in Eq. (3.44), with  $\tilde{V}_T$  defined in Eq. (3.40). Although a similar phenomenon has previously been observed in the literature for certain calculations pertaining to certain string models, in this paper we have formulated an appropriate definition for  $\tilde{V}_T$  which respects the full symmetries of the theory (such as modular invariance and  $T$ -duality), and then proceeded to prove the resulting scaling rule in full generality. Indeed, our proof holds even in cases without supersymmetry and regardless of the spacetime background, and assumes only that our amplitudes are finite. Our second  $T$ -volume scaling rule, by contrast, generalizes the first  $T$ -volume scaling rule to cases in which our theory contains certain massless states that cause our overall amplitude to have a logarithmic divergence. We nevertheless see that similar results apply.

Even beyond their phenomenological implications, the new supertrace constraints we have found could have far-reaching implications pertaining to the possible structure of string theory itself. Indeed one might reasonably ask if these constraints are so tight that the only solution is that all of the supertraces vanish level-by-level in a pairwise fashion. This would then be a cancellation between the contribution from a given state and that from a corresponding partner of opposite spin—i.e., a cancellation between states and their superpartners. This would then imply that the only consistent four-dimensional theories are those that can decompactify exclusively to higher-dimensional theories which are supersymmetric. Moreover, if the relevant compactification is untwisted, this would in turn imply that the original four-dimensional theory would have to be supersymmetric as well.

However, it is not true that the only way to solve these constraints is through boson/fermion pairings. Indeed, there exists an entire *landscape* of tachyon-free nonsupersymmetric heterotic string models [40,41], and each of these furnishes us with an explicit example of a modular-invariant spectrum within which such constraints are satisfied without boson/fermion pairings. It is not difficult to understand why such solutions exist. Within any of these string models the spectrum exhibits exponentially growing degeneracies of states. Moreover, each of these states has its own helicity and gauge charges. As a result, even within the bounds of modular invariance, the spectrum has far more ways of arranging itself and its charges than there are constraints on it. Indeed, the existence of an entire landscape of such strings provides direct verification of this fact. Of course, a separate question concerns the *stability* of such tachyon-free nonsupersymmetric strings. Addressing this question is beyond the scope of this work, although we

note that there exist nonsupersymmetric tachyon-free string models in four dimensions in which various instabilities can be exponentially suppressed [29].

We close with six important remarks. The first of these concerns the generality of our results with respect to our choice of regulator. By implicitly distinguishing which states are to be considered dynamical and which are to be considered heavy and therefore “integrated out,” the regulator helps to establish an appropriate energy scale  $\mu$  for an effective field theory derived from the string. However, as we have stated throughout this paper, our conclusions are independent of the specific form of the regulator  $\mathcal{G}_\rho(a)$ . This is ultimately due to the fact that our results stem directly from the modular invariance of the regulator, and any suitable regulator must be modular invariant. Indeed, the precise definition of energy scale cannot affect the constraints that must be satisfied in order to cancel the accumulated contributions from the infinite towers of string states. Likewise, our  $T$ -volume defined in Eq. (3.40) and the scaling rules which result from this definition are independent of the regulator as well.

Second, we note that the new constraints that we have found are also independent of the compactification radius. This in turn implies that the states which collectively correspond to a  $\delta > 2$  decompactification direction can *never* contribute to running, regardless of the size of the compactification volume. This remains the case even when the compactification volumes are of order the string scale, provided that the theory does not develop a tachyon at some radius. Indeed, so long as the theory is merely *capable* having a large-volume decompactification, the constraints that correspond to every possible decompactification direction must all simultaneously apply in the four-dimensional theory. As a result, our nonrenormalization theorems hold even beyond the framework of large-volume compactifications, and have a generality that matches that of their supersymmetric counterparts. Indeed, as discussed at the end of Sec. IV B, our results suggest that nonsupersymmetric theories seem to be governed by the same sorts of nonrenormalization theorems as apply to supersymmetric theories, and that UV/IR mixing appears to play many of the same roles as supersymmetry. This observation is worthy of further exploration.

Third, we note that our conclusion that the running of gauge couplings exhibits a fixed-point behavior at high scales bears a superficial similarity to the results of Ref. [42]. In Ref. [42], a purely field-theoretic analysis was performed—based on the results of Refs. [35,36]—in which it was shown that although the accumulating effects of Kaluza-Klein states would generically convert a four-dimensional logarithmic gauge-coupling running into a higher-dimensional power-law running, this power-law growth could occasionally be cancelled in certain theories by the effects of switching from the true four-dimensional gauge coupling to an effective loop expansion parameter

which is sensitive to the number of Kaluza-Klein levels that have been crossed at a given scale  $\mu$ . This would then result in an apparent UV fixed point for the higher-dimensional theory. However, despite this similarity, there is ultimately no connection between these results: in this paper we are considering the actual gauge couplings, not the effective loop-expansion parameter; we are not performing a field-theoretic analysis involving only Kaluza-Klein states but rather a fully modular-invariant analysis involving not only Kaluza-Klein states but also winding states; and our nonrenormalization theorem is completely general, serving as a core feature of *all* tachyon-free modular-invariant theories and emerging directly from the UV/IR-mixed modular invariance of the theory. Thus, despite these superficial similarities, the underlying physics is completely different.

Fourth, and somewhat connected to the previous point, we note that it is important to carefully consider the physical meaning of renormalization in the ultraenergetic regime with energies exceeding the string scale  $M_s$ . In particular the existence of an invariance under the scale-duality transformation  $\mu \rightarrow M_s^2/\mu$  and the corresponding “dual” behavior that we have described for  $\mu > M_s$  may seem paradoxical. Indeed, in the common “textbook” definition of renormalization, the energy scale  $\mu$  corresponds in some sense to the deep-Euclidean momenta of asymptotic states that are taking part in a scattering experiment. However, the very notion of preparing such highly energetic states as the momenta approach the Planck scale is fraught with difficulty due to gravitational back-reaction. By contrast, in this paper, our modular-invariant prescription for renormalization has allowed us to find well-behaved dual perturbative behavior at energy scales  $\mu \gg M_s$ .

To understand the meaning of this alternative behavior, we first note that below the scale  $M_s$  our definition of renormalization gives behavior that is broadly aligned with the conventional one (and with that of the several other definitions of renormalization that one could use). In other words, in both definitions of renormalization, one can think of energy scales as below  $M_s$  as corresponding to the momenta of scattering momentum modes. It is only above  $M_s$  that the two renormalization prescriptions appear to diverge. This apparently discrepant behavior arises because above the scale  $M_s$  the scaling behavior of effective couplings can be very different in different processes, such that there is no universal definition of renormalization. This difference between the two definitions of renormalization in the energy regime beyond  $M_s$  arises because the energy scales are effectively for asymptotic scattering states that are in fact not the same. The energy scale in the textbook definition of renormalization implies that one is always scattering ultrahigh momentum modes. By contrast in the present paper the energy scales beyond  $M_s$  correspond to processes that involve the scattering of ultramassive dual winding modes. This is unavoidably the case because the

modular-invariant prescription of renormalization that we have been using implies that only these particular processes enjoy a dual perturbative description with dual energy scales  $\mu < M_s$ . This discrepancy between renormalization definitions implies that great care is required when invoking the concept of renormalization at scales above  $M_s$ , since our entirely perturbative modular-invariant description of renormalization effectively works to automatically exclude this region by instead mapping ultrahigh stringy energies to a dual regime. Indeed, this situation is analogous to what happens in the case of  $T$ -duality, wherein one must be careful to distinguish between the different dual pictures when attempting to make any measurement of compactification volume [43]. Thus it would be very interesting to further explore definitions of scaling and renormalization in the region above  $M_s$ .

Fifth, throughout this paper, we have stressed that our results apply only to modular-invariant theories. Certainly perturbative Type II and heterotic strings fall into this class. However, as the string coupling increases, we know that these theories cease to be perturbative and become more accurately described in terms of dual theories such as M-theory or theories of Type I strings. However, M-theory is not governed by modular invariance. Likewise, although Type I strings have closed-string sectors (corresponding to the torus amplitudes) in which modular invariance continues to apply, such strings also contain twisted and open-string sectors (corresponding to the Klein-bottle, cylinder, and Möbius-strip amplitudes) in which modular invariance no longer strictly applies and in which only certain orientifold projections of modular invariance continue to apply. It would therefore be interesting to understand how the nonrenormalization theorem and supertrace identities we have derived in this paper are deformed or “loosened” as the string coupling grows beyond the perturbative regime. In this context, we remark that many of the properties that come from modular invariance (such as misaligned supersymmetry) apparently nevertheless have echoes within Type I strings [44–46]. It would be interesting to examine whether this is true for our nonrenormalization theorem as well.

Finally, we note that our work may have important connections to the recent swampland program which seeks to ascertain the limits of viability of four-dimensional theories as low-energy approximations to UV-complete theories of quantum gravity [47,48]. In particular, decompactification limits have played a central role in swampland discussions of the so-called “distance conjecture” (see, e.g., Refs. [48–74]). Our theorem clearly has relevance to this question. Indeed, while previous work in this area has mostly developed along generic lines, utilizing traditional concepts of energy scales, mass splittings, and relations to the cosmological constant, it is clear from our discussion that these concepts and relations may be highly modified when UV/IR mixing is properly taken into account. Indeed, as we have seen, a generic theory will not be able to

accommodate a decompactification limit unless it already possesses some very special and nongeneric properties. The ramifications of these ideas for the distance conjecture and for the swampland program as a whole will be discussed in future work.

## ACKNOWLEDGMENTS

The research activities of S. A. A were supported by the STFC Grant No. ST/P001246/1. The research activities of K. R. D. were supported in part by the U.S. Department of Energy under Grant No. DE-FG02-13ER41976/DE-SC0009913, and also by the U.S. National Science Foundation through its employee IR/D program. The opinions and conclusions expressed herein are those of the authors, and do not represent any funding agencies.

## APPENDIX: AMPLITUDES WITH $\bar{E}_2$ FACTORS

In this appendix we provide a prescription for evaluating the supertraces of operators whose definitions include factors of the Eisenstein function  $\bar{E}_2$ . The details behind this prescription are given in Ref. [17].

Toward this end, let us assume that we have an operator of the form

$$\mathcal{X} = \mathcal{A} + \bar{E}_2 \mathcal{B} \quad (\text{A1})$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are functions of  $\tau_2$ . Let us also assume that  $\mathcal{A}$  and  $\mathcal{B}$  have separate  $\tau_2$ -expansions of their own:

$$\begin{aligned} \mathcal{A} &= \mathbb{A}_0 + \tau_2 \mathbb{A}_1 + \tau_2^2 \mathbb{A}_2 \\ \mathcal{B} &= \mathbb{B}_0 + \tau_2 \mathbb{B}_1 + \tau_2^2 \mathbb{B}_2. \end{aligned} \quad (\text{A2})$$

Our goal is to evaluate  $\text{Str} \mathcal{X}$ , or more general expressions such as  $\text{Str}[\mathcal{X}f(M)]$  where  $f(M)$  is some function of the mass  $M$  of the states across the string spectrum.

To do this, let us recall that  $E_2(\tau)$  has its own  $q$ -expansion

$$E_2(\tau) = \sum_{r=0}^{\infty} \chi_r q^r \quad (\text{A3})$$

where the coefficients are given by

$$\chi_r = \begin{cases} 1 & r = 0 \\ -24\sigma(r) & r > 0 \end{cases} \quad (\text{A4})$$

with  $\sigma(r)$  denoting the sum-of-divisors function  $\sigma(r) \equiv \sum_{d|r} d$ . For example, we find  $\sigma(r) = 1, 3, 4, 7, \dots$  for  $r = 1, 2, 3, 4, \dots$

Given this definition for the coefficients  $\chi_r$ , we now define what we shall call an “ $E$ -entwined” supertrace, to be denoted  $\text{Str}_E$ . This is given in terms of the  $\chi_r$ -coefficients as

$$\text{Str}_E \mathcal{X} \equiv \sum_{r=0}^{\infty} \chi_r \text{Str}^{(r)} \mathcal{X} \quad (\text{A5})$$

where  $\text{Str}^{(r)} \mathcal{X}$  denotes the *shifted* supertrace over all the states in the theory that are “level-mismatched” by  $r$  units:

$$\text{Str}^{(r)} X \equiv \lim_{\tau_2 \rightarrow 0} \sum_n a_{n-r,n} X_{n-r,n} e^{-\pi \alpha' M_L^2 \tau_2}. \quad (\text{A6})$$

In Eq. (A6), as in Eq. (2.3), the quantity  $M_L$  denotes the left-moving contribution to the total mass of the state, i.e.,  $\alpha' M_L^2 = 4n$ , while  $a_{n-r,n}$  is the net number of (bosonic minus fermionic) states which have

$$\alpha'(M_L^2 - M_R^2) = 4r. \quad (\text{A7})$$

Our claim, then, is that  $\text{Str}[\mathcal{X}f(M)]$  can be evaluated by formally replacing [17]

$$\text{Str}[\mathcal{X}f(M)] \longrightarrow \text{Str}[\mathcal{A}f(M)] + \text{Str}_E[\mathcal{B}f(M_L)]. \quad (\text{A8})$$

Equivalently, for each component  $\mathbb{X}_\ell$ , we may replace

$$\text{Str}[\mathbb{X}_\ell f(M)] \longrightarrow \text{Str}[\mathbb{A}_i f(M)] + \text{Str}_E[\mathbb{B}_i f(M_L)]. \quad (\text{A9})$$

This procedure applies to all of our expressions in Eq. (2.39).

With these replacements, our theorem then continues to apply exactly as described in the main text.

- 
- [1] K. R. Dienes, *Nucl. Phys.* **B429**, 533 (1994).  
[2] K. R. Dienes, M. Moshe, and R. C. Myers, *Phys. Rev. Lett.* **74**, 4767 (1995).  
[3] K. R. Dienes, *Nucl. Phys.* **B611**, 146 (2001).  
[4] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, *Phys. Rev. Lett.* **82**, 4971 (1999).  
[5] S. A. Abel, J. Jaeckel, V. V. Khoze, and A. Ringwald, *J. High Energy Phys.* **01** (2006) 105.  
[6] S. A. Abel, J. Jaeckel, V. V. Khoze, and A. Ringwald, *J. High Energy Phys.* **09** (2006) 074.  
[7] J. Bramante and E. Gould, *Phys. Rev. D* **101**, 084022 (2020).  
[8] N. Craig and S. Koren, *J. High Energy Phys.* **03** (2020) 037.  
[9] N. Blinov and P. Draper, *Phys. Rev. D* **104**, 076024 (2021).  
[10] P. Charalambous, S. Dubovsky, and M. M. Ivanov, *Phys. Rev. Lett.* **127**, 101101 (2021).  
[11] N. Craig, I. G. Garcia, A. Vainshtein, and Z. Zhang, *J. High Energy Phys.* **05** (2022) 079.  
[12] A. Castellano, A. Herráez, and L. E. Ibáñez, *J. High Energy Phys.* **08** (2022) 217.  
[13] S. Abel and K. R. Dienes, *Phys. Rev. D* **104**, 126032 (2021).  
[14] N. Arkani-Hamed and K. Harigaya, *J. High Energy Phys.* **09** (2021) 025.  
[15] L. Freidel, J. Kowalski-Glikman, R. G. Leigh, and D. Minic, *Phys. Rev. D* **107**, 126016 (2023).  
[16] P. Berglund, T. Hübsch, and D. Minic, *Int. J. Mod. Phys. D* **32**, 2330002 (2023).  
[17] S. Abel, K. R. Dienes, and L. A. Nutricati, *Phys. Rev. D* **107**, 126019 (2023).  
[18] N. Craig, *Eur. Phys. J. C* **83**, 825 (2023).  
[19] P. Berglund, L. Freidel, T. Hübsch, J. Kowalski-Glikman, R. G. Leigh, D. Mattingly, and D. Minic, *arXiv:2202.06890*.  
[20] R. Rankin, *Proc. Cambridge Philos. Soc.* **35**, 351 (1939).  
[21] R. Rankin, *Proc. Cambridge Philos. Soc.* **35**, 357 (1939).  
[22] A. Selberg, *Arch. Math. Naturvid.* **43**, 47 (1940).  
[23] D. Zagier, *J. Fac. Sci. Tokyo* **28**, 415 (1982).  
[24] D. Kutasov and N. Seiberg, *Nucl. Phys.* **B358**, 600 (1991).  
[25] G. W. Moore, *Nucl. Phys.* **B293**, 139 (1987); **B299**, 847(E) (1988).  
[26] K. R. Dienes, *Phys. Rev. Lett.* **65**, 1979 (1990).  
[27] K. R. Dienes, *Phys. Rev. D* **42**, 2004 (1990).  
[28] I. Kani and C. Vafa, *Commun. Math. Phys.* **130**, 529 (1990).  
[29] S. Abel, K. R. Dienes, and E. Mavroudi, *Phys. Rev. D* **91**, 126014 (2015).  
[30] R. Rohm, *Nucl. Phys.* **B237**, 553 (1984).  
[31] R. Hagedorn, *Nuovo Cimento Suppl.* **3**, 147 (1965).  
[32] J. J. Atick and E. Witten, *Nucl. Phys.* **B310**, 291 (1988).  
[33] K. R. Dienes, M. Lennek, and M. Sharma, *Phys. Rev. D* **86**, 066007 (2012).  
[34] S. A. Abel, K. R. Dienes, and L. A. Nutricati (to be published).  
[35] K. R. Dienes, E. Dudas, and T. Gherghetta, *Phys. Lett. B* **436**, 55 (1998).  
[36] K. R. Dienes, E. Dudas, and T. Gherghetta, *Nucl. Phys.* **B537**, 47 (1999).  
[37] L. J. Dixon, V. Kaplunovsky, and J. Louis, *Nucl. Phys.* **B355**, 649 (1991).  
[38] V. S. Kaplunovsky, *Nucl. Phys.* **B307**, 145 (1988); **B382**, 436(E) (1992).  
[39] C. Angelantonj, I. Florakis, and B. Pioline, *Commun. Num. Theor. Phys.* **6**, 159 (2012).  
[40] K. R. Dienes, *Phys. Rev. D* **73**, 106010 (2006).  
[41] K. R. Dienes, M. Lennek, D. Senechal, and V. Wasnik, *Phys. Rev. D* **75**, 126005 (2007).  
[42] K. R. Dienes, E. Dudas, and T. Gherghetta, *Phys. Rev. Lett.* **91**, 061601 (2003).  
[43] R. H. Brandenberger and C. Vafa, *Nucl. Phys.* **B316**, 391 (1989).  
[44] N. Cribiori, S. Parameswaran, F. Tonioni, and T. Wrase, *J. High Energy Phys.* **04** (2021) 099.  
[45] C. Angelantonj, I. Florakis, and G. Leone, *J. High Energy Phys.* **06** (2023) 174.  
[46] G. Leone, *J. High Energy Phys.* **11** (2023) 066.  
[47] C. Vafa, *arXiv:hep-th/0509212*.

- [48] H. Ooguri and C. Vafa, *Nucl. Phys.* **B766**, 21 (2007).
- [49] F. Baume and E. Palti, *J. High Energy Phys.* **08** (2016) 043.
- [50] D. Klaewer and E. Palti, *J. High Energy Phys.* **01** (2017) 088.
- [51] R. Blumenhagen, I. Valenzuela, and F. Wolf, *J. High Energy Phys.* **07** (2017) 145.
- [52] R. Blumenhagen, D. Kläwer, L. Schlechter, and F. Wolf, *J. High Energy Phys.* **06** (2018) 052.
- [53] T. W. Grimm, E. Palti, and I. Valenzuela, *J. High Energy Phys.* **08** (2018) 143.
- [54] H. Ooguri, E. Palti, G. Shiu, and C. Vafa, *Phys. Lett. B* **788**, 180 (2019).
- [55] T. W. Grimm, C. Li, and E. Palti, *J. High Energy Phys.* **03** (2019) 016.
- [56] P. Corvilain, T. W. Grimm, and I. Valenzuela, *J. High Energy Phys.* **08** (2019) 075.
- [57] S.-J. Lee, W. Lerche, and T. Weigand, *J. High Energy Phys.* **02** (2022) 096.
- [58] S.-J. Lee, W. Lerche, and T. Weigand, *J. High Energy Phys.* **02** (2022) 190.
- [59] R. Blumenhagen, D. Kläwer, and L. Schlechter, *J. High Energy Phys.* **05** (2019) 152.
- [60] F. Marchesano and M. Wiesner, *J. High Energy Phys.* **08** (2019) 088.
- [61] A. Font, A. Herráez, and L. E. Ibáñez, *J. High Energy Phys.* **08** (2019) 044.
- [62] T. W. Grimm and D. Van De Heisteeg, *J. High Energy Phys.* **03** (2020) 020.
- [63] D. Lüst, E. Palti, and C. Vafa, *Phys. Lett. B* **797**, 134867 (2019).
- [64] F. Baume, F. Marchesano, and M. Wiesner, *J. High Energy Phys.* **04** (2020) 174.
- [65] N. Gendler and I. Valenzuela, *J. High Energy Phys.* **01** (2021) 176.
- [66] F. Baume and J. Calderón Infante, *J. High Energy Phys.* **08** (2021) 057.
- [67] E. Perlmutter, L. Rastelli, C. Vafa, and I. Valenzuela, *J. High Energy Phys.* **10** (2021) 070.
- [68] B. Bastian, T. W. Grimm, and D. van de Heisteeg, *J. High Energy Phys.* **06** (2021) 162.
- [69] D. Klaewer, S.-J. Lee, T. Weigand, and M. Wiesner, *J. High Energy Phys.* **03** (2021) 252.
- [70] J. Calderón-Infante, A. M. Uranga, and I. Valenzuela, *J. High Energy Phys.* **03** (2021) 299.
- [71] M. Graña and A. Herráez, *Universe* **7**, 273 (2021).
- [72] T. Rudelius, *J. High Energy Phys.* **08** (2021) 041.
- [73] M. Montero, C. Vafa, and I. Valenzuela, *J. High Energy Phys.* **02** (2023) 022.
- [74] C. Aoufia, I. Basile, and G. Leone, [arXiv:2405.03683](https://arxiv.org/abs/2405.03683).